Abstract

A recent literature documents a secular increase in the sales-weighted average markup in the United States, a phenomenon that was driven by large and productive firms at the top of the profit distribution. Using rich administrative balance-sheet data, this paper documents the behaviour of markups in Spain before, during, and in the aftermath of the Great Recession. We show that markups rose during the financial crisis. Unlike in the U.S., these dynamics were driven by small and unproductive firms following an increase in their average costs due to a high and increasing share of fixed costs, especially overhead labour expenses. Despite the rise in markups, profit rates fell, especially amongst smaller firms.

Keywords: Markups, Market Power, Average Costs, Labour Market, Firm Size.

JEL codes: D2, D4, E2, E3, J3, L1

*A previous version of this paper was circulated under the title Raising Markups to Survive: Small Spanish Firms during the Great Recession. For helpful comments and discussions, we would like to thank the co-editor, the associate editor and two anonymous referees, as well as Oscar Arce, Olympia Bover, Eduardo Gutiérrez, Enrique Moral-Benito, Jorge Padilla, Roberto Ramos, Alberto Urtasun, Javier Vallés, and seminar participants at Bank of Spain, Bank of Portugal, the 44th Simposio de Análisis Económico (Alicante), Funcas, the OECD's Global Forum on Productivity 2020 Annual Conference and the Monetary Policy Committee (MPC) of the European Central Bank. The opinions and analyses are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem. The authors declare that they have no conflicting financial interests to report. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Any errors are our own.

†Banco de España, DG Economics and Research, Calle Alcalá 48, 28014 Madrid, Spain. Email: pilargarcia@bde.es.
‡Banco de España. Email: aitor.lacuesta@bde.es.
§Banco de España. Corresponding author. Email: pau.roldan@bde.es.
1 Introduction

The recent debate on the evolution of market power in the United States has reached a certain consensus. De Loecker et al. (2020) document a secular increase in the sales-weighted average markup, from about 20% in 1980 to nearly 70% in 2014. This rise was mostly driven by a within-industry reallocation of economic activity from low- to high-markup firms, but also by increasing markups at the firm level, keeping the distribution fixed. Furthermore, these dynamics were accompanied by a rise in firm profitability, suggesting that these top-markup firms increased their market power and contributed to the rise in industry concentration. These findings, along with other macroeconomic dynamics seen during the same period, are consistent with the secular decline in the labour share (the share of value added accounted for by aggregate labour expenditures), observed since the 1980s and especially during the 2000s (see e.g. Karabarbounis and Neiman (2013)). Autor et al. (2017, 2020) find, in particular, that most of the aggregate labour share decline in the United States can be attributed to a within-industry reallocation of activity among heterogeneous firms toward those with low and declining labour shares.\textsuperscript{1} Their explanation relies on the emergence of superstar firms, namely a small number of firms which are able to harness a large share of the market through winner-takes-most dynamics.\textsuperscript{2}

There is less consensus, however, regarding whether similar patterns have emerged in other parts of the world and, particularly, in Europe. On the one hand, Autor et al. (2020) find that the evolution of the aggregate labour share in Europe is qualitatively similar to that of the United States. On the other hand, Gutiérrez and Philippon (2018) argue that patterns of rising concentration and rising profits rates are not as evident in Europe as they are in the U.S.

This paper contributes to the debate by showing new evidence on the evolution of markups, profit rates, and concentration for Spain (2004-2017). Using detailed administrative balance-sheet data on the quasi-universe of Spanish firms (including private and public, as well as small and large,

\textsuperscript{1}Though different explanations for the fall in the labour share have been proposed in the literature, there is a tight conceptual relationship between the labour share and markups. When there is imperfect competition in output markets, the pass-through of cost shocks to output prices is incomplete because output demand is not perfectly elastic, leading to a partial adjustment in variable input choices and an increase in the price markup. In the United States, where labour inputs face low adjustment costs and may be considered variable, a negative correlation between firm-level markups and labour shares is therefore to be expected.

\textsuperscript{2}Kehrig and Vincent (2021) similarly find that the decline in the U.S. labour share is driven by large firms.
firms), we find that markups increased by about 13% between 2004 and 2009, and have been on a slight decreasing trend since. Similar to the United States, these dynamics are led by firms at the top of the markup distribution. However, in stark contrast to the U.S. experience, the evolution of markups was driven primarily by firms at the lower end of the productivity distribution, whilst firms at the top experienced small or even negligible changes in markups. During the crisis years, the aggregate dynamics were explained by a large group of small firms increasing their markup, rather than by the reallocation of activity across firms within industries. In the recovery period, small firms continued to increase their markups in spite of losing market share to larger firms.

We argue that the key explanation behind these new findings lies in the cost structure of firms, and in particular in the allocation of expenditures between fixed and variable inputs. Making use of the granularity of our data, we show that firms of different sizes differ in their allocation of expenditures between different types of costs. For smaller firms, fixed inputs (such as quasi-fixed operating expenses and labour costs from workers with open-ended contracts) represent a higher share of their total sales. Moreover, smaller firms experience stronger procyclical variation in their variable input use. Therefore, when estimating markups against such variable inputs (materials, in our baseline estimation), smaller firms appear to charge higher markups. Interestingly, these small, unproductive, high-markup firms also have higher labour shares on average, which we argue is due to labour inputs being subject to high adjustment costs in Spain. Indeed, we show that the average markup behaves very differently if it is estimated with respect to labour: it falls sharply in the 2004-2009 period (between 13 and 16%, depending on the weighting scheme), and rises back to near 2004 levels by the end of our sample.

A key contribution of our study is, therefore, to show that in countries with frictional input markets, the behavior of markups may be reflecting not only changes in market power but also, and perhaps more importantly, underlying changes in the allocation of expenditures across different types of costs. All in all, our results are in stark contrast to the empirical findings in the U.S., where

\[\text{In the production-function approach that we use for estimation, markups are a wedge between the variable input cost share of sales and the output elasticity to this input (see Section 2 for details). Therefore, a lower material share by small firms translates into a higher markup, other things equal.}\]

\[\text{The Spanish labour market is characterized by a two-tier system, with the co-existence of open-ended contracts with large termination costs, and fixed-term contracts of short duration (e.g. Bentolila et al. (2019)). In fact, the share of fixed-term contracts in Spain (34% at its peak in 2006, according to data from the Labor Force Survey provided by the Spanish Statistical Office) is amongst the highest in all of Continental Europe.}\]
the evolution of markups more likely reflected a rise of market power by large and very productive firms rather than an effort by firms to restructure the composition of costs in response to the cycle. Our main takeaway is, therefore, that the evolution of markups may not reflect aggregate changes in the competitive structure of markets, but rather an idiosyncratic response by firms to macroeconomic conditions.

The Spanish experience is of particular interest in the European context because of at least three reasons. First, it offers an insightful case study for the evolution of markups for firms at the low end of the productivity distribution. Indeed, according to a Banco de España (2019) annual report, the average productivity of Spanish firms is lower than that of French, German and Italian firms, regardless of their size, in part because Spanish firms tend to have lower levels of human capital and technology. Moreover, this gap relative to neighboring European economies is overwhelmingly high for smaller firms. Indeed, Spanish firms on average underscore the productivity level of their European counterparts by around 40 percentage points.

Second, the available firm-level data in Spain, from the Spanish Commercial Registry, is of high quality both in terms of coverage and, most importantly for our purposes, in terms of balance-sheet information. In particular, our dataset covers 80% of all limited responsibility firms, includes publicly listed as well as private firms, and has a particularly good coverage of small firms. The great level of detail in balance-sheet information allows us to separate out variable inputs (such as materials and workers with fixed-term contracts) from fixed inputs (such as other operating expenses and workers with open-ended contracts). This is a notable improvement in terms of data quality relative to the existing markup estimation literature using the production-side approach of Hall (1988) and De Loecker and Warzynski (2012). An important contribution of our paper is to show how such granularity in terms of balance-sheet information allows us to uncover previously unexplored sources of markup heterogeneity coming from differences in the input use across firms that differ in size and productivity.

Third, the Spanish case is of particular interest because of the high degree of frictions of the Spanish labour market, which allows us to interpret labour as an input with high adjustment costs. Yet, labour costs make up for a sizable fraction of the typical Spanish firm’s balance sheet, and exhibit interesting cyclical properties, with sharp increases during recessions that are especially pronounced.
amongst smaller firms.\footnote{There are also large differences from the worker’s point of view. For instance, during the crisis period, workers with open-ended contracts experienced almost no loss in terms of earnings (see Anghel et al. (2018)).} Indeed, the dual market structure of labour in Spain puts constraints on the ability for firms to freely adjust their workforce. For instance, larger Spanish firms make disproportionate use of internal flexibility mechanisms to reduce labour-related costs within the firm, such as the renegotiation of wages and working time schedules, the worker’s geographic location, or reductions in monetary and non-monetary payments over the worker’s salary (see Bertola et al. (2010)).\footnote{These measures were introduced by the various labour market reforms that took place in Spain during our sample period.} Smaller firms, by contrast, tend to make less use of these mechanisms and, by and large, tend to adjust their labour costs via the extensive margin. All in all, given the complex nature of labour markets in Spain, a complete picture of the behaviour of firm-level markups requires the analysis of the evolution of input shares over time separately for each input in the production function of firms.

**Related Literature** Our paper contributes to a literature estimating markups for different countries. Analyzing whether their findings could be generalized at a global scale, De Loecker and Eeckhout (2018) find a worldwide upward trend for markups. Within Europe, Spain and Portugal exhibit the lowest markup increase, and well below the observed growth in the United States over the same period. Similarly, Autor et al. (2017) show that the decline of the labour share in Spain is one of the least pronounced in Europe, and much lower than that in the United States.

There is less consensus regarding the evolution of profit rates in Europe compared to the United States. De Loecker and Eeckhout (2018) calculate that both economic regions have comparable growth rates for this variable, whereas Gutiérrez and Philippon (2018) argue that patterns of rising concentration and rising profits rates are not visible in Europe. One possible explanation that is suggested by the authors is that while U.S. markets were more competitive than Europe’s until the 1990s, this trend has reversed in the two last decades, perhaps due to differences in antitrust enforcement and product market regulations having become more aggressive in Europe than in the U.S. in recent years. Moreover, Díez et al. (2021) argue that competition has declined at a global scale, though markups have increased only modestly, while Díez et al. (2018) find a stronger increase, though only among publicly listed firms. We contribute to these studies by providing
evidence based on very granular data covering all types of firms (large and small, private and public) and exploiting the rich cost structure of firm balance sheets.7

Our paper also contributes to a strand of the literature that studies the cyclical properties of markups. Generally speaking, we confirm the insight made by De Loecker and Warzynski (2012) that markup measurement is highly sensitive to the production input that the markup is estimated on. In the context of the Spanish economy, estimating markups using the inverse of the labour share would lead to the conclusion that markups are procyclical, as remarked by Domowitz et al. (1986) and Nekarda and Ramey (2020) for the United States. In the case of Spain, this is because labour can be thought of as a quasi-fixed input whose expenditure share increases in recessions, especially amongst small firms. A similar conclusion is reached then markups are estimated based on intermediate inputs that are unrelated to the scale of production (e.g. advertising, as in Hall (2012)). On the other hand, considering only materials, as in Anderson et al. (2018), or taking into account that a fraction of labour costs are related to overhead workers, as in Galí et al. (2007) or Nekarda and Ramey (2020), would lead to countercyclical variation in markups.8

In the case of Spain, large firms exhibit lower labour cost shares from permanent workers and a lower share of fixed operating expenses compared to small firms. When comparing the evolution over the crisis years of these items of the balance sheet, we find that these inputs appear to be more rigid for smaller firms: the evolution of fixed inputs’ shares is positively correlated with our preferred markup estimates (computed using materials), making markups more countercyclical for smaller firms.

Outline The remainder of the paper proceeds as follows. Section 2 describes the data and presents the methodological framework. Section 3.1 shows the evolution and several decompositions of markups in Spain. In Section 3.2, we offer potential explanations for the observed dynamics based on changes in the composition of costs across firms during this period. Section 3.3 analyses the effects on profit rates and market concentration. Section 4 offers some concluding remarks.

7Even though our data does not go back farther than 2004, the cyclical behaviour of markups across firm size that we find is consistent with different patterns of growth observed in Spain (and in other Southern European countries) during the previous expansionary period (1995-2007). As documented in García-Santana et al. (2019) and Gopinath et al. (2017), for example, misallocation of resources increased during the expansionary period, with unproductive firms gaining market share.

8There is an active debate in the literature regarding the sign and magnitude of the cyclicality of markups, including Bils (1987), Hall (1988), Stroebel and Vavra (2019), Bils et al. (2018), and Burstein et al. (2020).
2 Data and Methodology

Data Our data cover the period 2004-2017, and come from an unbalanced panel of administrative and confidential firm-level data from the Spanish Commercial Registry (Registro Mercantil Central). The dataset presents an excellent coverage of the market economy (see Almunia et al. (2018)), spanning the last few years of the expansionary period (2004-2007), the crisis years (2008-2013), and the first few years of the recovery (2014-2017). After our cleaning filters, the final dataset has approximately 3.8M firm-year observations, with about 300,000 firms per year on average. Sector information of the firm is at the 4-digit NACE Rev. 2 level.

Importantly, the information in the database allows us to identify various elements of the cost structure of firms: (i) material inputs, or purchases of inputs that affect the level of production; (ii) other operating expenses (OOPE), composed of a heterogeneous group of costs (e.g. services provided by other independent firms, renting, transportation, utilities, insurance, professional services, R&amp;D or marketing), and which are not directly related to the level of production; (iii) labour costs, computed as compensation to the employees of the firm, with additional information on the number of workers employed with fixed-term contract versus those with open-ended ones; and (iv) financial expenses, including interest payments, depreciation, and amortization. This level of cost disaggregation will allow us to argue that the evolution of markups in Spain is, to a large degree, driven by changes in the composition of balance sheets during the cycle.

Estimating Markups Firm-level markups are estimated using the production-approach method, developed by Hall (1988) and De Loecker and Warzynski (2012), which has gained popularity in the literature. This method has the advantage that it does not require postulating a demand system or any assumption on the market’s competitive structure. Instead, as the method is production-based, it only relies on firm-level revenue and input expenditure information from balance-sheet data. However, it also has some important caveats. First, the method presumes that the cost-minimization

---

9We drop observations with missing or zero sales, employees, materials, or fixed assets, and drop outliers from the labour and material share distribution, since these inputs will be used in the production function estimation step. Moreover, we focus on industries that have at least ten firms per year, and use 2-digit-level value-added deflators for materials, sales, and fixed assets, which we take from the Spanish National Accounts.

10This item of the balance sheet is analogous to the Selling, General and Administrative (SG&amp;A) Expenses in other datasets, such as Compustat.
problem is static, as the firm’s choice is over a variable input which can be costlessly adjusted between periods. In reality, however, there likely exist dynamic pricing considerations related to either input or output demand inertia that potentially affect the markup, e.g. in the form of customer capital or via firm expectations about the business cycle. Second, though the method does not need to impose any model of output demand and competition, it does impose some structure on input markets. In particular, it assumes that the firm takes input prices as given when making their input choices. Though this does not preclude these prices to incorporate some (unobserved) markup from the input supplier, this may nevertheless distort our markup estimates. In Appendix B we show how to proxy for these imperfections in labour input markets, following recent work by Morlacco (2020) and Mertens (2020).

Methodology  In an economy with \( i = 1, \ldots, N \) cost-minimizing firms each year \( t \), firm \( i \) is assumed to operate the gross-output production function \( Q(\Omega_{it}, V_{it}, F_{it}) \), where \( \Omega_{it} \) is firm-specific productivity, which is observed by the firm when choosing its inputs but not by the econometrician; \( F_{it} \) is a vector of fixed and quasi-fixed inputs, including capital and other inputs with high adjustment costs such as labour employed under open-ended contracts, and other operating expenses such as intangible investments, outsourcing expenditures or utilities; and \( V_{it} = (V^1_{it}, \ldots, V^J_{it}) \) is a vector of intermediate variable inputs with no adjustment costs, such as materials and temporary employment.

The first-order condition of minimizing costs with respect to variable input \( V \in V \) is:

\[
\frac{\partial Q_{it}}{\partial V_{it}} \cdot \frac{V_{it}}{Q_{it}} = \frac{1}{\Lambda_{it}} \cdot \frac{P^V_{it} V_{it}}{Q_{it}}
\]

where \( \Lambda_{it} \geq 0 \) is the Lagrange multiplier. The left-hand side of this equation is the elasticity of output to variable input \( V \), denoted \( \theta^V_{it} \equiv \frac{\partial Q_{it}}{\partial V_{it}} \frac{V_{it}}{Q_{it}} \). As \( \Lambda_{it} \) is the shadow value of the objective function as the constraint \( Q_{it} \geq Q(\cdot) \) is relaxed, it proxies the marginal cost of the firm. Therefore, we may define the firm’s markup (the ratio of the output’s price \( p^Q_{it} \) to the marginal cost) as \( \mu_{it} \equiv \frac{p^Q_{it}}{\Lambda_{it}} \).

Using the optimality condition, we then obtain a formula for the markup:

\[
\mu_{it} = \theta^V_{it} \cdot V S^{-1}_{it}
\]

\(^{11}\)See also Rubens (2021) for a related approach.
where \( VS_{it} \equiv \frac{P^V_{it}}{P^Q_{it}} \) is the “\( V \)-share”, i.e. the share of sales that are accounted for by expenditures on variable input \( V \). In words, the firm-level markup is inversely proportional to the firm’s variable input's share of sales, with a factor of proportionality given by the elasticity of output to that input. Our goal is then to estimate \( \mu_{it} \) for every Spanish firm and year from 2004 to 2017 by using equation (1), and study how estimates differ depending on the choice of \( V \).

For this, we need to estimate (i) the output elasticity of the variable input to production (\( \theta^V_{it} \)); and (ii) the input’s share of sales (\( VS_{it} \)). Note that \( VS_{it} \) can be directly obtained from our data, by taking the ratio of the total expenditures on input \( V \) to sales for each firm. To estimate \( \theta^V_{it} \), we need to posit a functional form for the production function, \( Q(\cdot) \). The only assumptions on the production function are: (i) unobserved productivity \( \Omega_{it} \) is Hicks-neutral; (ii) technology parameters are time-invariant and common across producers within the same sector.\(^{12}\) That is, \( Q(\Omega_{it}, V_{it}, F_{it}) = \Omega_{it} \tilde{Q}(V_{it}, F_{it}; \beta) \), where \( \beta \) is the vector of industry-specific technology parameters. To exploit as much variation as possible, we estimate elasticities at the most disaggregated level available in the data, namely 4-digit industries.\(^{13}\) In the data, we rely on log-sales deflated using 2-digit sectoral value-added deflators, \( y_{it} \equiv \ln Y_{it} \), and assume that they equal desired output (\( q_{it} \equiv \ln Q_{it} \)) plus a term \( \epsilon_{it} \) capturing unanticipated productivity shocks and possible measurement error (both of which unobserved by the firm when choosing inputs), so that \( y_{it} = q_{it} + \epsilon_{it} \). Therefore, the specification is:

\[
y_{it} = \omega_{it} + \tilde{q}(v_{it}, f_{it}; \beta) + \epsilon_{it} \tag{2}
\]

where lower-case letters denote logged variables. Here, the left-hand side is given directly by the data and, on the right-hand side, the parameter vector \( \beta \) must be estimated. Estimating (2) directly by OLS could potentially suffer from simultaneity bias (if unobserved productivity shocks in \( \epsilon_{it} \) are correlated with input choices), serial correlation bias (if the productivity \( \omega_{it} \) has correlated effects), and selection bias (if, over time, sample selection occurs among exiting firms). To deal with these issues, we proceed using the Olley and Pakes (1996) two-stage approach. First, we express unobserved productivity as an (unknown) function of the firm’s state variables:

\(^{12}\)Note that there is no requirement of constant returns to scale, so the method can accommodate the existence of overhead costs, which will be crucial in our subsequent analysis.

\(^{13}\)Moreover, we do not consider changes of technology as a consequence of the crisis, as this would have been problematic given the limited sample size at this level of disaggregation.
\[
\omega_{it} = h_t(v_{it}, f_{it})
\]

(3)

where \(h_t\) is a non-parametric function (e.g. a polynomial) of variable inputs, capital, labour, possibly other fixed inputs, and time dummies. By positing (3), we are presuming that current input use responds to current productivity shocks, but lagged input values do not. Under this assumption, output \(q_{it}\) is now proxied by:

\[
\phi(v_{it}, f_{it}; \beta) \equiv h_t(v_{it}, f_{it}) + \tilde{q}(v_{it}, f_{it}; \beta)
\]

(4)

Then, we run OLS on \(y_{it} = \phi(v_{it}, f_{it}; \beta) + \epsilon_{it}\). Using the estimate \(\hat{\beta}\), we can predict expected output, \(\hat{\phi}_{it} \equiv \phi(v_{it}, f_{it}; \hat{\beta})\), and the residual, \(\hat{\epsilon}_{it} = y_{it} - \hat{\phi}_{it}\). Now, we can compute productivity for any \(\beta\) via equation (4), \(\omega_{it}(\beta) \equiv \hat{\phi}_{it} - \tilde{q}(v_{it}, f_{it}; \beta)\). To make progress, we assume that \(\omega_{it}\) follows an AR(1): \(\omega_{it} = \rho \omega_{i,t-1} + \xi_{it}\). That is, by equation (3), we presume that lagged input use is correlated with current input use only through the innovations in the productivity process, or, in other words, through serial correlation in input market conditions (by virtue of equation (3)). Next, projecting \(\omega_{it}(\beta)\) on its lag \(\omega_{i,t-1}(\beta)\), we recover the innovations as \(\xi_{it}(\beta) \equiv \omega_{it}(\beta) - \rho \omega_{i,t-1}(\beta)\). To obtain our final estimate for \(\beta\), we use the moment condition:

\[
\mathbb{E} \left[ \xi_{it}(\beta) \begin{pmatrix} v_{i,t-1} \\ f_t \end{pmatrix} \right] = 0
\]

From here, we can estimate \(\beta\) by GMM. Using the estimate \(\hat{\beta}_{GMM}\), we can directly calculate the markup as \(\hat{\mu}_{it} = \hat{\theta}_V^t \cdot V^{-1}\). Here, \(\hat{\theta}_V^t\) is the output elasticity implied by the GMM estimates.

3 Empirical Results

3.1 Markup Analysis

Our firm-level markups use materials as the baseline variable input of reference. We do this because, given the rigidities in the Spanish labour market, this is the input in our balance-sheet data with arguably the least adjustment costs in production. To implement the estimation procedure, we
use a Cobb-Douglas production function of the type:

\[ q_{it} = \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it} \]  

(5)

implying that the estimated markup is given by \( \mu_{it} = \hat{\beta}_m \cdot MS_{it}^{-1} \), where \( MS_{it} \equiv \frac{MaterialCosts_{it}}{Sales_{it}} \) is the share of sales accounted for by expenditures in material inputs (or, in short, the “material share”), and \( \hat{\beta}_m \) is the GMM estimate of the elasticity of output to materials, specific to each 4-digit sector in the data.\(^{14}\) We then average markups using firm-level sales shares, as follows:

\[ \mu_t = \frac{\sum_{i=1}^{N_t} \sigma_{it} \mu_{it}}{\sum_{i=1}^{N_t} \sigma_{it}}, \quad \text{where} \quad \sigma_{it} = \frac{Sales_{it}}{\sum_{i=1}^{N_t} Sales_{it}} \]  

(6)

Figure 1 shows the evolution of the sales-weighted markup, normalized to one at the base year (2004). Markups were relatively constant between 2004 and 2007 and sharply increased between 2008 and 2009. The overall rise of markups between 2004 and 2009 was of around 13ppt, a comparable increase to what was observed in the U.S. (see De Loecker et al. (2020)). Since 2009, though, markups have been on a steady decrease, by some 5ppt until 2017.

![Figure 1: Sales-weighted average markup in Spain, normalized to one for the year 2004.](image)

Figure 2 shows the evolution of markups by different branches of activity, compared to that of the whole economy from Figure 1. Markups in nearly all sectors picked up between 2007 and 2009. However, sectoral heterogeneity arises in terms of the changes thereafter. Markups in manufacturing

\(^{14}\)To avoid outliers, we winsorize the top and bottom 5% of observations in all of our markup estimates.
and in some service sectors fell down to levels similar to those of the initial period. On the contrary, a structural increase in markups during the 2009-2017 period is observed in Supplies, Construction, and Real Estate. The moderate decline of markups in most service sectors is in contrast to the patterns found in the U.S. by De Loecker et al. (2020).

![Graph of sales-weighted average markup, by major sector of activity (dashed line) and in the overall economy (solid line). Markups normalized to one for the year 2004.](image)

**Robustness** According to equation (6), the behaviour of the average markup can be explained by changes in (i) the material share, (ii) the output elasticity, and (iii) the sales weights. We will argue that the most relevant dynamics arise from (i). To rule out (ii) and (iii) as important drivers, we conduct the following robustness exercises.\(^\text{15}\)

\(^{15}\)Additionally, in Figure 8 we will argue that the change in markup is mostly driven by firm-level changes in markups rather than by the reallocation of sales shares within industries over time.
First, Figure A1 in the Appendix compares our baseline results to a Translog production function, which allows for time-varying elasticities of output to the variable inputs. The results are not affected in any significant way by this assumption, and the markup still increases sharply in the crisis. Therefore, our markup dynamics are likely not driven by changes in output elasticities.

Second, as weighting by sales shares is somewhat contentious in the literature (see e.g. Edmond et al. (2021)), Figure A2 presents the cost-weighted average markup based on the five input costs that we have available in our data, of which three are quasi-fixed (overall labour costs, cost from permanent workers, and cost from other operating expenditures, the part of intermediate inputs that does not change with firm output) and two are quasi-variable (materials and costs from temporary workers). When weighting with variable costs (materials and temporary labour), the increase in markups in the 2004-2009 period is higher, whilst the increase in the average markup is very close to the one obtained with sales weights if we use any of the three permanent inputs as cost weights. Yet, qualitatively, the markup continues to rise in the recession, and decline thereafter.

Other Inputs  In the Appendix, we show results for alternative specifications in which markups are estimated relative to some of the other inputs that have been traditionally considered in the literature. Using overall labour expenses (Figure A3) instead of materials leads to a completely different evolution of the average markup. Indeed, markups relative to labour dropped sharply between 2008 and 2009 and steadily recovered after 2010. However, computing markups using labour expenses related to temporary workers (Figure A4) leads to a qualitatively very similar picture as in the baseline estimation using materials, whereby markups increased during the crisis, and decreased shortly after. As it will become clear in Section 3.2, a large share of the adjustment of labour following the recession was made via the firing of workers with temporary contracts, which account for about a third of total compensation of employees in Spain during this period. This explains that the evolution of the labour share of those who hold a permanent contract as a fraction

\[ \mu_{it} = \hat{\beta}_l L_{Si}^{-1} , \]

where \( \hat{\beta}_l \) is the GMM estimate of \( \beta_l \) and \( L_{Si} \) denotes the labour share of firm \( i \) (i.e. the ratio of the firm’s wage bill to its sales). To produce Figure A4, we assume instead a more general production function that includes temporary (\( l_T^{it} \)) as well as permanent (\( l_P^{it} \)) workers, so that \( q_{it} = \beta_T^T l_T^{it} + \beta_P^T l_P^{it} + \beta_m m_{it} + \beta_k k_{it} + \omega_{it} + \varepsilon_{it} \). The markup in this case is \( \mu_{it} = \hat{\beta}_l^T T L_{Si}^{-1} \), where \( T L_{Si} \) is the labour share from temporary workers (i.e. total costs from temporary labour divided by firm sales). Note that the first specification is a special case of the second one, imposing that the two types of labour are perfect substitutes (\( \beta_T^T = \beta_P^T = \beta_l \)).
of total costs is very different from that of materials and temporary workers.

**Cross-Sectional Decompositions** The remainder of this section presents several decompositions that help understand which type of firms set higher markups during this period, and which increased their markups the most.

As a first pass, Figure 3 shows the relationship between firm-level markups and market shares (within a 4-digit industry), labour shares (labour expenses over sales), and productivity (computed as sales per worker), respectively. The upper panel shows that firms that are large relative to other firms in their 4-digit industries set relatively lower markups, whereas the lower-left panel shows that firms with high labour shares set higher markups, and the lower-right panel shows that lower-markup firms are on average more productive. This is at odds with the hypothesis that more productive firms are able to set higher markups because they operate with lower marginal costs. In this case, in contrast, small and unproductive firms appear to use markups to partially counterbalance their disadvantage in terms of average costs.

Figure 4 presents the evolution of markups by different percentiles. The evolution of markups is highly asymmetric, being strongly countercyclical at the upper end of the distribution. The result is similar to the one found in De Loecker et al. (2020), with the exception that in their case the right tail gathered a group of large and very productive firms. In contrast, in the Spanish data this group is composed of small and unproductive firms who, given their cyclical behaviour of variable inputs, drive the increase in the overall average markup. Markups in the remaining deciles steadily increased and reached a 5 percentage-point rise in 2017 compared to 2004.

An alternative way of showing that small firms are setting higher markups is by decomposing the weighted average markup into two terms: a simple average of firm level markups, \( \bar{\mu}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \mu_{it} \), and a term that is proportional to the covariance between relative firm size and the relative level of markups:

---

\(^{17}\)To construct this figure, we first rank firms by markup decile, then compute the median market share, labour share and productivity within each decile, and average out across all years in our sample.

\(^{18}\)To construct this plot, first we rank firms by their raw markup, and then we compute the cumulative sum of sales shares across ranked firms, \( \hat{\sigma}_{it} \equiv \sum_{j=1}^{i} \sigma_{jt} \). The percentile \( p \in \{25, 50, 75, 90\} \) of the markup distribution in year \( t \) is then defined by \( \mu_{it}^p \equiv \max \{ \mu_{it} : \hat{\sigma}_{it} \leq p/100 \} \).
Figure 3: Median market share (within a 4-digit industry), labour share and productivity, by decile of the markup distribution, on average across years 2004-2017.

Figure 4: Evolution of the sales-weighted markup distribution, by percentile. All markup percentiles are normalized to one for the year 2004.
Figure 5 shows the result of this decomposition. We observe that the unweighted markup grew more during the period 2004-2009 than the weighted one, reflecting that markups were increasing more rapidly among small firms during this period. According to the right-hand side panel of the figure, the initial covariance between the size of the firm and markups was negative, suggesting that small firms were already setting higher markups at the start of the sample. Moreover, during those years, the covariance became even more negative, indicating that small firms were increasing their markups relative to large firms. After 2012, though, the gap between small and large firms starts to close again.

The countercyclical pattern of markups for the unweighted average holds true at the sectoral level as well. To see this, we perform a similar decomposition by sector. Notice we can compute our baseline measure of sales-weighted markups by taking sectoral sales weights as follows: $\mu_t = \sum_s \sigma_{st} \mu_{st}$, where $\sigma_{st} \equiv \frac{\sum_{j \in s} P_{jt} Q_{jt}}{\sum_i P_{it} Q_{it}}$ is the sales share of sector $s$ in the economy, and $\mu_{st} \equiv \sum_{i \in s} \sigma_{it} \mu_{it}$. Then, we may decompose the weighted sector average between an unconditional average across sectors and a covariance term between the sector weight and the sector markup:

$$\mu_t = \sum_i \sigma_{it} \mu_{it} = \bar{\mu}_t + \sum_i (\sigma_{it} - \bar{\sigma}_t) (\mu_{it} - \bar{\mu}_t)$$

Covariance size vs. markup

Figure 5: Left: Sales-weighted markup ($\mu_t$) and unweighted markup ($\bar{\mu}_t$). Both markup measures are normalized to one for the year 2004. Right: Covariance term between relative firm size and relative markup.
\[
\mu_t = \bar{\mu}_t + \sum_s (\sigma_{st} - \bar{\sigma}_s) (\mu_{st} - \bar{\mu}_t)
\]

Covariance between sector size and markup

Figure 6 shows the result of this decomposition when we define a sector \( s \) as a 2-digit or a 4-digit industry. We find, similar as with the firm-level analysis, that those sectors with lower markups on average represent a larger share of overall value added. During the period between 2008 and 2017, the covariance became less negative, however, reflecting that those sectors with higher markups were gaining weight during the recovery period.

Figure 6: Left: Sales-weighted markup (\( \mu_t \)), and average sectoral markup (\( \bar{\mu}_t \)) when a sector is defined as a 2-digit (dashed line) or a 4-digit (dotted line) industry. All markups are normalized to one for the year 2004. Right: Covariance term between relative sectoral size and relative markup.

Figure 7 shows this decomposition by major sector of activity. The speed at which the unweighted average of markups decreased after 2010 was different across sectors, being very rapid in Manufacturing and relatively slow in Supplies and some of the service sectors. The covariance term is negative in all sectors (see Figure A5 in the Appendix), but there exist substantial differences in its dynamics. In Supplies and Construction, for instance, the gap in markups between small and large firms shrinks starting in 2010, signaling changes in the composition of these industries. In other sectors, by contrast, such as Retail, IT or Transportation, the gap widens between firm size and markup (relative to the sectoral average), with smaller firms within a sector being the main drivers of the average markup in the aftermath of the recession.
Figure 7: Sales-weighted (solid line) and unweighted (dashed line) average markup, by major sector of activity. Markups normalized to one for the year 2004.
**Time-Series Decomposition** The evidence presented thus far shows that, cross-sectionally, small firms have higher markups in our data. But which margins contribute to the change in the average markup over time? Is the increase in markups during the crisis explained by small firms increasing their markups, or is it explained by demographic and reallocate changes within industries? To answer these questions, we decompose the year-on-year change in the sales-weighted average markup $\Delta \mu_t \equiv \mu_t - \mu_{t-1}$ as follows:

$$
\Delta \mu_t = \sum_i \sigma_{i,t-1} \Delta \mu_{it} + \sum_i \left( \mu_{i,t-1} \Delta \sigma_{i,t} + \Delta \mu_{i,t} \Delta \sigma_{i,t} \right) + \sum_{i \in E_t} \sigma_{i,t} (\mu_{it} - \mu_{t-1}) - \sum_{i \in X_t} \sigma_{i,t-1} (\mu_{i,t-1} - \mu_{t-1})
$$

(7)

where $E_t$ is the set of firms active in $t$ and inactive in $t - 1$ (i.e. entering firms), and $X_t$ is the set of firms active in $t - 1$ and inactive in $t$ (i.e. exiting firms). This equation decomposes the change in markups into the contribution of new entrants, firms that exit the market, and changed amongst incumbents. In turn, this latter component is split between the change in firm-level markups keeping their initial weight in the sales distribution fixed (within effect), and the change in markups attributed to changes in those shares (reallocation effect). Figure 8 shows the contribution of each margin.\(^{19}\) Note that the contribution of the composition of firms in the industry (entry and exit components) is only minor compared to the reallocation and within components. Interestingly, the reallocation component, which is responsible for most of the markup increase in the U.S. during the Great Recession (De Loecker et al. (2020)), contributes negatively in Spain, indicating that higher-markup firms have decreasing shares of the market. By contrast, the within component contributes very positively, meaning that the increase in the average markup is dominated by changes in firm-level markups, and not by a redistribution of market shares in the industry. In fact, the increase in the average markup during the crisis years is solely driven by small firms increasing

\(^{19}\)To construct this figure, we run different counterfactual scenarios by setting the initial markup to the level in 2004, and then cumulatively adding the changes of each component in equation (7) assuming that other components do not contribute. For example, the dashed black line shows the evolution of the average markup that results from assuming that all changes come from the within component, keeping the other changes at zero. The other lines are constructed similarly.
markups, for given market shares.

![Firm decomposition of change in Markup](image)

Figure 8: Decomposition of the change in markups in within, reallocation, entry and exit components.

Therefore, the bulk of the change in the aggregate markup is accounted for by the firm-level markup dynamics of incumbent firm, with entrants and exiters playing only a minor role. This is in spite of the crisis years being characterized by a sharply declining inflow and increasing outflow of firms from the economy. Figure 9 plots the entry and exit rates in the Spanish economy during this period, constructed as follows:

\[
\text{EntryRate}_t = \frac{B_t}{\frac{A_t + A_{t-1}}{2}} \quad \text{and} \quad \text{ExitRate}_t = \frac{D_t}{\frac{A_t + A_{t-1}}{2}}
\]

where \(B_t\) and \(D_t\) denote the number of births and deaths in year \(t\), and \(A_t\) is the number of active firms. We observe an increase in exit rates in 2009 from slightly above 4% to 8%. During the remainder of the crisis and the recovery periods, exit rates have been steadily declining at a low pace.

\[20\text{The data used in this figure comes from the Central Business Register (Directorio Central de Empresas, or DIRCE) constructed by the Instituto Nacional de Estadística, the Spanish Statistical Office, an administrative dataset containing the yearly record of all firm births and deaths.}\]
3.2 Structure of Costs

So far, we have argued that markups rose in the crisis years because small and unproductive firms increased their markups, with little changes in the composition of industries. We now argue that, rather than reflecting changes in market power, these dynamics are the result of the reallocation of expenditures by firm between variable and fixed inputs over the business cycle.

In our data, variable costs are composed of a fraction of intermediate goods (materials) and a small fraction of labour costs including those workers holding temporary contracts, whereas fixed costs include some additional operating expenses, the bulk of labour costs, and financial expenses. To understand how firms changed their cost expenditures as a share of sales, we first look at the evolution of sales for different firms. Figure 10 plots the evolution of the distribution of log-turnover over time for different percentiles of the distribution.\footnote{Turnover is deflated using 2-digit sectoral value-added deflators.} The picture is clear: firms at the bottom of the sales distribution suffered a disproportionally large decrease in sales relative to firms at the top. Moreover, the reduction in sales was longer-lived for smaller firms.

Figure 11 shows that there exist sizable differences in Spain in the cost structure of firms by size, especially regarding the contribution of materials and labour expenses. In particular, the weighted average material share represents around 60% of sales, whereas this ratio is close to 45% for the
unweighted average. This means that large firms spend more than small ones in materials relative to their sales, explaining their lower markups.\footnote{One interesting hypothesis is that differences in the material share between small and large firms may be explained by differences in input inventory policy. However, exploring input inventory data, we have observed that material inventories tend to experience much larger shifts over time for large firms than for smaller firms. Thus, if anything, large firms seem to rely more on adjusting costs via inventory accumulation than smaller firms. Data available upon request.} By contrast, there are no such differences in other operating expenses (second panel on the top row), this being the balance-sheet item that, together with materials, adds up to overall expenditures in intermediate goods of the firm. Indeed, the weighted ratio is close to 15%, with the unweighted average fluctuating between 20 and 25%.

Regarding labour expenses, small firms spend more as a share of sales than large firms do, and this difference counterbalances the relatively higher material share faced by small firms. One important question is whether this labour input can be considered a fixed or a variable input. Our dataset allows us to decompose labour costs between workers with different types of contract. This is important in Spain because open-ended contracts enjoy a higher level of legal protection from the side of the worker, so we may consider these as relatively fixed inputs compared to workers in fixed-term positions. We find that, for all firms, most of the labour cost is related to workers with open-ended contracts, with the sales-weighted average cost related to fixed-term contracts accounting for around 4% of revenues and the unweighted average for 8%. Hence, most of the differences among firms here stem from those labour expenses associated with the share of labour that has higher adjustment costs. Interestingly, though, as a fraction of total labour costs, the

Figure 10: Evolution of the distribution of turnover (in logs, deflated with 2-digit value-added deflators), by percentiles. All percentiles are normalized to one for the year 2004.
costs related to permanent workers are slightly higher for larger firms. Finally, financial expenses (including interest payments, depreciation and amortization) account for between 1.5% and 2.5% of total revenue, regardless of the size of the firm. Adding up materials and temporary employees as variable inputs, large firms have a higher share of variable inputs with respect to small firms. Therefore, small firms operate with higher fixed costs than large firms (in both general operating expenses and especially in terms of labour costs), which explains the negative covariance between firm size and markups.

The evolution of these shares over time offers some interesting additional insights. During the pre-crisis expansionary period (between 2004 and 2007), all ratios for both the weighted and the unweighted averages remain constant, the only exceptions being the unweighted average of the material share, which decreases 2ppt during these years. Between 2007 and 2013, the crisis period, there was a decrease in the material share, which was somewhat lower for the weighted average (2ppt) than for the unweighted average (3ppt). In the first two years of the crisis, all firms decreased their material shares between 3ppt and 4ppt. A potential explanation for these dynamics is that material inventories are being depleted temporarily during recessions, thereby decreasing the share of revenues accounted for by expenses in new materials. To check this hypothesis, we analysed data (available upon request) on material inventories for all sectors of the economy, and observed that changes in

\[23\]
some 2ppt, while the unweighted average remained constant. This evolution is a key driver of the evolution of markups for small firms (indeed, recall that the measured markup is inversely proportional to the material share).

Other general expenses and labour costs exhibit similar dynamics. The crisis years saw an increase in labour costs for all firms, but the increase was again higher for smaller firms. In this case, the timing also shows some differences between large and small firms. Large firms initiated a process of decreasing labour costs in 2009, whereas small firms delayed this process to 2013. Most of the rise in labour costs is due to the increase in the open-ended component. After 2010, large firms initiated a reduction of the permanent component that was not matched by smaller firms until 2013. This is consistent with evidence found by Bertola et al. (2010), which argues for the difficulties of small firms to use internal flexibility measures to reduce labour costs along the intensive margin. The Annual Labour Cost Survey 2013 of the Spanish Ministry of Labour shows that 23.8% (28.7%) of firms between 5 and 9 workers (between 10 and 49) used internal flexibility measures to reduce labour costs within the firm (such as the renegotiation of wages and time schedules, the geographic location of workers or reductions in monetary and non-monetary payments over the salary), whereas this percentage peaks at 44% for firms between 50 and 249 workers, at 47% for firms between 250 and 499, and at 57% for firms over 500 workers. By contrast, the share of costs that can be attributed to temporary workers decreased relative to sales, more so for small firms, between 2007 and 2010.

Finally, financial expenses increased slightly between 2007 and 2009, initiating a process of continuous reduction that accelerated between 2011 and 2012. After 2013, during the recovery years, both the weighted and the unweighted average ratio of materials and temporary labour costs to turnover recovered slightly, more so for the latter. On the other hand, general expenses decreased and so did the labour share. Financial expenses over sales fell more than 1ppt for all firms.

Summing up, it appears as though small and unproductive firms, which experienced problems to contain the sharp rise in their fixed costs (related to labour inputs and other general expenses),

---

inventories are relatively larger in construction, real estate and manufacturing. In construction, firms replenished their input inventories during the crisis. In real estate, we observe large swings in inventories, but they are not cyclical and therefore may not fully explain the observed dynamics of markups in this sector. Finally, in manufacturing, we see a continued increase in inventory accumulation, which may partly explain why the measured markup exhibits a low rate of growth in this sector (recall Figure 2) compared to others.
increased markups the most. To provide a more formal check of this pattern, we run the following regression:

$$\ln \left( \frac{\mu_{is,t}}{\mu_{is,t_0}} \right) = \beta_0 + \beta_1 \Delta \alpha_{is,t_0} + \beta_2 \text{FinStress}_{is,t_0} + \lambda_j + \epsilon_{ist}$$

In this equation, the dependent variable is the log difference in the firm-level markups of firm $i$ from sector $s$ between years $t_0$ and $t$, $\Delta \alpha_{is,t_0}$ is the log change in the share of total sales of the firm represented by fixed inputs (including financial expenses, labour costs and other general expenses), and $\text{FinStress}_{is,t_0}$ is a dummy variable indicating if the firm is under financial distress (identified as a point in time in which financial expenses are greater than profits). We explore three sub-periods: pre-crisis (2004-2007), crisis (2007-2013), and recovery (2013-2016). To capture time-invariant industry-specific factors that are common across firms, we control for industry fixed effects, $\lambda_j$.

Table 1: Relationship between markup, average costs and financial stress. All standard errors (in parentheses) clustered by 4-digit industry. Significance: $* = 10\%$, $** = 5\%$, $*** = 1\%$.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \alpha_{is,t_0}$</td>
<td>0.6298***</td>
<td>0.6033***</td>
<td>0.5793***</td>
</tr>
<tr>
<td></td>
<td>(0.0276)</td>
<td>(0.0242)</td>
<td>(0.0225)</td>
</tr>
<tr>
<td>$\text{FinStress}_{ist}$</td>
<td>0.0326***</td>
<td>0.0434***</td>
<td>0.0015***</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0088)</td>
<td>(0.0054)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.23</td>
<td>0.24</td>
<td>0.20</td>
</tr>
<tr>
<td># Obs.</td>
<td>173,147</td>
<td>149,076</td>
<td>237,179</td>
</tr>
</tbody>
</table>

Table 1 shows the results. We find a statistically significant relationship between the change in the markup and the change in the share of fixed inputs at the firm level for all three sub-periods of analysis: those firms who increased the share of fixed inputs over sales the most saw, on average, an increase in their markups. Moreover, markup changes are also positively correlated with financial stress, consistent with the idea that those firms that are most financially constrained also increased their markups in order to reduce the effects of increased costs onto their profits.

3.3 Profit Rates and Concentration

Common wisdom in economics associates markups with profitability and market power. We have argued thus far that both the cross-sectional and the time-series patterns of markups in Spain
can be largely explained by the allocation of expenditures across different types of cost. The picture would not be complete, however, without analysing the behavior of profit rates and concentration measures during this period, a question to which we turn next.

To understand the connection between markups, costs, and profit rates, it becomes useful to recall the following basic accounting identity:

\[ \frac{\pi_{it}}{Sales_{it}} = 1 - \frac{1}{\mu_{it}} \frac{AC_{it}}{MC_{it}} \]

This firm-level identity shows that the profit rate (the ratio of profits to sales) is increasing in the markup and inversely proportional to the ratio between average costs (AC) and marginal costs (MC). Average and marginal costs would coincide in the absence of economies of scale (i.e. with constant returns to scale in production) and with no fixed costs. As seen before, however, this is not the case in Spain. Therefore, to understand the connection between profitability and markups, one must understand the behaviour of both variable and fixed costs.

Figure 12 shows the evolution of the profit rate, where profits are defined as earnings before interest, tax, depreciation and amortization (EBITDA). The first thing one should notice is that, despite having higher markups, small firms retain a smaller share of their revenues as profit: while the sales-weighted average profit rate is close to 10% for the period before the crisis, it is only 5% for the unweighted average. During the crisis, profit rates decreased for both weighted and unweighted averages, but the drop was much more pronounced for the unweighted average, indicating that smaller firms decreased their profit margins by relatively more. In the period after 2012, there was a slight recovery of profits, although not strong enough to reach the levels observed in 2007 (9% for the weighted average and 3% for the unweighted). These findings are in line with the evidence on profit rates presented in Gutiérrez and Philippon (2018). Overall, the results suggest that the growth in markups in Spain was not accompanied by a rise in the profitability of firms, suggesting that the evolution of markups may indeed be unrelated to market power and instead connected with the internal structure of costs of firms.

A similar picture emerges at the sector level as well (Figure 13). Indeed, both unweighted and weighted averages in virtually all sectors face a procyclical movement in profit rates, with smaller firms suffering deeper and more persistent losses. In Supplies, there is an increase in the weighted
Figure 12: *Left:* Sales-weighted and unweighted profit rates, computed at the firm level as the ratio of earnings before interest, tax, depreciation and amortization (EBITDA) to sales. *Right:* Covariance term between relative firm size and relative profit rate.

Figure 13: Sales-weighted (solid line) and unweighted (dashed line) average profit rate, by major sector of activity. Profit rates are computed as the ratio of earnings before interest, tax, depreciation and amortization (EBITDA) to sales.
average profit rate from 20% in 2007 to more than 25% in 2017. Profit rates experienced a notable recovery in some sectors such as Transportation, Accomodation and Real Estate. On the other hand, profit rates in IT and Construction continued to decline even after the recession was over.

To close, we look at industry concentration. In our analysis above we have shown how different firms within a particular market (defined by a 4-digit industry) faced different costs, set different markups and obtained different profits during the recession years. If we were to understand markups as reflecting market power, those small and unproductive firms that were setting higher markups between 2007 and 2013 should have gained market share in their sector, which should lead to a reduction in industry concentration. To test this explanation, Figure 14 plots the sales share of the 10 largest firms within a 4-digit industry, averaged across all industries. The evidence reflects a rise in concentration between 2007 and 2013, followed by a period of decreasing concentration during the recovery. Figure 15 shows that this occurred across most branches of activity as well. Therefore, the evidence suggests that, rather than reflecting changes in the dynamics of market power within industries, our observed markup dynamics are more likely the reflection of the re-structuring of the portfolio of intermediate input expenditures by firms.
Figure 15: Unweighted average concentration ratios of 10 largest firms, by major sector of activity (solid lines), together with average concentration ratio at the economy-wide level (dashed lines).
4 Conclusion

Using administrative balance-sheet data containing the quasi-universe of firms and all input types for the period 2004-2017 in Spain, this paper documents that markups were higher and increased during the Great Recession for small firms compared to their larger competitors in the same sector. We argue that, rather than reflecting market power, these differences are explained by the allocation of costs across different types of inputs. In particular, we show that the share of materials and temporary labour costs of small firms is substantially lower and declined more dramatically amongst small and unproductive firms during the crisis years. Simultaneously, the labour share from expenses in permanent workers and overhead costs over sales were much higher and increased more for small firms during this period. Moreover, the rise in the aggregate markup during the crisis cannot explained by a reallocation of activity within industries, but rather by smaller firms increasing their markups. In the aftermath of the recession, small firms continued to raise their markups, in spite of continuously losing market share.

Through the lens of our analysis, the different results found for the United States by De Loecker et al. (2020) and Gutiérrez and Philippon (2017), among others, may be rationalized for certain Southern European economies like Spain by firms being smaller on average and having to operate in more frictional input markets. These circumstances might, in turn, make average costs less resilient to recessionary shocks, while making it harder for firms to benefit from technological progress and globalization. Further exploring the differences between firms of different countries along these dimensions remains an interesting avenue for future research.
References


Anghel, Brindusa, Henrique Basso, Olympia Bover, José María Casado, Laura Hospido, Mario Izquierdo, Ivan A. Kataryniuk, Aitor Lacuesta, José Manuel Montero, and Elena Vozmediano, 2018, Income, consumption and wealth inequality in Spain, SERIES 9, 351 – 387.


Banco de España, 2019, The Spanish economy and the more uncertain global enviroment. Recent developments, outlook and challenges, Banco de España Annual Report 1.

Bentolila, Samuel, Juan J. Dolado, and Juan F. Jimeno, 2019, Dual Labour Markets Revisited, IZA Discussion Paper No. 12126.


31


Markups and Cost Structure: Small Spanish Firms during the Great Recession
by Pilar Garcia-Perea, Aitor Lacuesta and Pau Roldan-Blanco

Appendix Materials

A Additional Figures

Figure A1: Sales-weighted markup under different production function specifications. The translog production function is: 
\[ q_{it}^{\text{TL}} = \beta_{l} l_{it} + \beta_{m} m_{it} + \beta_{k} k_{it} + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{lm} l_{it} m_{it} + \beta_{lk} l_{it} k_{it} + \beta_{mk} m_{it} k_{it} + \omega_{it} + \epsilon_{it} \]
and thus the markup is 
\[ \mu_{it}^{\text{TL}} = (\beta_{m} + 2\beta_{mm} m_{it} + \beta_{lm} l_{it} + \beta_{mk} m_{it}) \cdot MS_{it}^{-1} \]
where \( MS_{it} \equiv \frac{\text{Materials}_{it}}{\text{Sales}_{it}} \).
Figure A2: Cost-weighted average markup, under different cost weighting schemes: “Labour” means weights based on total labour costs; “Perm labour” means weights based on labour costs from permanent workers (i.e. workers employed under an open-ended contract); “Temp labour” means weights based on labour costs from temporary workers (i.e. workers employed under a fixed-term contract); “Materials” means weights based on material expenditures; “OOPE” means weights based on other operating expenditures.

Figure A3: Sales-weighted (solid line) and unweighted (dashed line) average markup relative to labour input, normalized to one for the year 2004. The production function is \( q_{it} = \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it} \), and the markup is computed as \( \mu_{it} = \hat{\beta}_l \cdot LS_{it}^{-1} \), where \( \hat{\beta}_l \) is the GMM estimate of \( \beta_l \) and \( LS_{it} \) denotes the labour share of firm \( i \).
Figure A4: Sales-weighted (solid line) and unweighted (dashed line) average markup relative to temporary labour input, normalized to one for the year 2004. The production function is $q_{it} = \beta^T l^T_{it} + \beta^P p^P_{it} + \beta^m m_{it} + \beta^k k_{it} + \omega_{it} + \epsilon_{it}$, and the markup is computed as $\mu_{it} = \hat{\beta}^T \cdot TLS^{-1}_{it}$, where $TLS_{it}$ is the labour share from temporary workers (i.e. total costs from temporary labour divided by firm sales).
Figure A5: Covariance term between relative firm size (within the sector) and relative firm-level markup, by major sector of activity.
B Computing Markups under Imperfect Input Markets

A key argument in this paper is that the Spanish economy is characterized by firms who differ in their use of variable and fixed inputs. We show that this circumstance is relevant to understand the dynamics of markups in Spain. One limitation of the production approach used to estimate markups in this paper, however, is that firms take input prices as given. In this appendix, we show a generalization of the method, based on insights by Mertens (2020) and Morlacco (2020), which allows us to quantify such input market frictions within the context of our markup estimation method.

To start, assume that the production function is, similar to our baseline specification, given by:

\[ Q_{it} = Q(\Omega_{it}, K_{it}, F_{it}, M_{it}) \]

where \( K_{it} \) is capital, \( F_{it} \) is an input obtained from a frictional input market, \( M_{it} \) is materials (or an otherwise variable input), and \( \Omega_{it} \) is Hick-neutral productivity. In this specification, \( F_{it} \) stands for any input with a quasi-fixed nature among the ones discussed in the paper, e.g. labour, other operating expenses, and otherwise quasi-fixed intermediate inputs that may be subject to market power on either the buyer or the seller side (e.g. long-term contracts between the firm and its suppliers, bargaining power on the side of workers, etc.).

Each firm now solves the cost-minimization problem:

\[
\min_{K_{it}, F_{it}, M_{it}} r_{it}K_{it} + p^F(F_{it})F_{it} + p^M_{it}M_{it} \quad \text{s.t.} \quad Q_{it} \geq Q(\Omega_{it}, K_{it}, F_{it}, M_{it})
\]  

(8)

where \( r_{it}, p^F(F_{it}) \) and \( p^M_{it} \) are the unit input costs of capital, quasi-fixed inputs and materials. Notice that the input price \( p^M_{it} \equiv p^F(F_{it}) \) is written as a function of \( F_{it} \), reflecting that frictions in this input’s market imply that the firm does not take prices as given. However, materials’ input prices are competitive and, for simplicity, so is the rental rate of capital. We shall continue to denote the output price by \( P_{Q_{it}} \).

The first-order condition with respect to material inputs implies:

\[ \mu_{it} = \theta^M_{it} \cdot (MS_{it})^{-1} \]  

(9)

where \( \theta^M_{it} \equiv \frac{\partial Q_{it}}{\partial M_{it}} \frac{M_{it}}{Q_{it}} \) is the output elasticity with respect to materials, \( MS_{it} \equiv \frac{p^M_{it}M_{it}}{P_{Q_{it}}Q_{it}} \) is the material share, and we have used that the Lagrange multiplier equals \( \Lambda_{it} = \frac{P_{Q_{it}}Q_{it}}{\mu_{it}} \), as \( \Lambda_{it} \) proxies the marginal cost of the firm. While \( \theta^M_{it} \) needs to be estimated, \( MS_{it} \) can be readily computed from our firm-level data.

The first-order condition relative to the quasi-fixed input, in contrast, implies:

\[ \mu_{it} = \gamma_{it} \cdot \theta^F_{it} \cdot (FS_{it})^{-1} \]  

(10)
where $\theta^F_{it} \equiv \frac{\partial Q_{it}}{\partial F_{it}} F_{it}$ is the output elasticity with respect to the quasi-fixed input, and $FS_{it} \equiv \frac{p^F_{it} F_{it}}{p^M_{it} M_{it}}$ is the input’s share. Comparing the two optimality conditions, we observe that the optimal choice for $F_{it}$ has an extra term $\gamma_{it}$ in it, defined as follows:

$$\gamma_{it} = \frac{\varepsilon^F_{it}}{1 + \varepsilon^F_{it}}, \quad \text{where} \quad \varepsilon^F_{it} = \frac{\partial p^F_{it} F_{it}}{\partial F_{it} p^F_{it}}$$

In this equation, $\varepsilon^F_{it}$ is the price-elasticity of the quasi-fixed input. In fact, notice that $\gamma_{it} = \frac{p^F_{it}}{MFC_{it}}$, where $MFC_{it} \equiv \Lambda_{it} \frac{\partial Q_{it}}{\partial L_{it}}$ is the marginal revenue cost of input $F$. If the input market were competitive, no wedge would exist, and the price and marginal cost would be equalized. However, because of input market imperfections, there is a wedge on input prices.

To calculate $\gamma_{it}$ in the data, we combine (9) and (10) to find:

$$\gamma_{it} = \frac{\theta^M_{it}}{\theta^F_{it}} \cdot \frac{p^F_{it} F_{it}}{p^M_{it} M_{it}}$$

(11)

It is easy to show that this formula for input market imperfections is independent of the side of the market from which the market power comes from (see e.g. Mertens (2020)), and we can therefore use it as a general proxy for the degree of input market frictions in the markup estimation.

Using these insights, Figure A6 shows the sales-weighted and unweighted average $\gamma$ wedges, by major sector of activity, and normalized to one for 2004. We find that, in all sectors, there is a negative correlation between firm size and the wedge. Since $\gamma$ is proportional to the ratio of labour costs to material costs, the negative correlation is consistent with the idea that smaller firms use a higher share of fixed inputs than of variable inputs. Moreover, this relative share increases during the crisis in all sectors, reflecting the reallocation of expenditures away from variable inputs (e.g. materials), which once again occurs in all sectors.
Figure A6: Sales-weighted (solid line) and unweighted (dashed line) average $\gamma$ wedge in labour inputs, by major sector of activity, normalized to one for the year 2004. The wedge is computed using equation (11).