The Macroeconomics of Market Power*

JAN DE LOECKER
University of Leuven†

JAN EECKHOUT
UPF Barcelona‡

October 29, 2025

Abstract

This review article surveys the literature on the macroeconomics of market power. Since 1980, various aggregate measures of market power, such as markups, profit rates and market capitalization, have increased, in the US and around the globe. We discuss the challenges of measuring economy-wide market power, as well as a few leading likely determinants of this change. We find that the rise of aggregate markups is due in large part to the reallocation of market shares, as well as the rise of the inverse expenditure share on variable inputs. Instead, output elasticities have not risen. We then analyze the macroeconomic causes and consequences of this process, and use a model to evaluate the potential underlying mechanisms. Key in the transmission mechanism is the underlying heterogeneity in the distribution of parameters. This passes through to heterogeneity in the distribution of outcomes such as output, the labor share, profits and welfare. We illustrate that focusing on aggregate measures of market power is misleading as the same aggregate markups are associated with different output and welfare, depending on how the markups and underlying parameters are distributed. We conclude with a discussion of the role of policy, in particular, how competition policy can be a tool not only for restoring economic efficiency, but also for redistribution.

Keywords. Market Power in the Aggregate Economy. Market Structure. Conduct. Technological Change. Reallocation. Endogenous Markups. Wage Stagnation. Labor Share. Passthrough. Business Dynamism. Profits. Stock Market Valuation.

*We are grateful for comments and feedback from Basile Grassi. Eeckhout acknowledges support from the ERC, Advanced grant 882499, from AEI (Barcelona School of Economics CEX2019-000915-S), and from PGC2018-096370-B-I00. Alessandro Fornasari provided superb research assistance. https://doi.org/10.1146/annurev-economics-100825-105527

†jan.deloecker@kuleuven.be

‡jan.eeckhout@upf.edu – ICREA-BSE-CREI

Contents

1 Introduction **2**

2 Economy-wide Market Power: Facts and Measurement **5**

2.1 Measuring Market Power in the Macroeconomy 5

2.2 The Components of Markups 9

2.3 Aggregate Markups and Heterogeneity 12

3 The Macroeconomic Causes and Consequences of Market Power **15**

3.1 Causes and Consequences 16

3.2 A Model 21

3.3 Quantification and Comparative Statics 24

3.4 Aggregate Outcomes when Aggregate Markups are Constant 25

4 Competition Policy as Redistribution **28**

5 Conclusion **31**

A Model **38**

A.1 Omitted Derivations 38

A.2 Solution Algorithm 39

B Additional Figures **41**

B.1 Measures of Market Power 41

B.2 Comparative Statics 46

C Markup Decomposition **52**

1 Introduction

Market power is a macroeconomic phenomenon. Market power is crucial for the transmission of interest rates and monetary policy, it determines optimal taxation and fiscal policy, and it affects inequality and social policy. Since 1980, there has been a marked rise in various measures of market power, chief among them margins, profit rates and market capitalization.[1] This raises several macroeconomic questions. How do we measure market power in the aggregate? How does market power affect the rest of the economy? What are the welfare implications of economy-wide market power?

Modern macroeconomics builds on individual behavior of firms and households, and studies the aggregate implications of joint behavior. In analyzing market power, the macroeconomic methodology builds on the micro approach developed in Industrial Organization. Firm and customer optimization decisions provide a framework for the measurement of markups, a measure of market power that gauges the pricing power of the firm. Firm-level production and cost data then enable the estimation of technologies and preferences on which those optimization decisions are based. This is in contrast with the tradition of modeling industries as a representative firm, subscribing to the notion that omitting firm heterogeneity is a useful tool to study the macroeconomic performance of countries and regions. This Harberger tradition dominated macroeconomic thinking for a substantial time, and consequently failed in obtaining crucial insights into the underlying drivers of aggregate outcomes, and therefore the associated policy responses. Most notably, we find that an increase in the dispersion of firm productivity is the most prominent driver behind the rise in market power.

There are, however, advantages from focusing on specific markets. In general, one can find more detailed data to estimate richer models of consumer preferences and firm production technology; we have better knowledge of market structure and conduct in specific markets; and competition policy often needs concrete answers about the impact of market power (e.g., mergers) in a given market. The macro approach to market power most often cannot address these pressing questions. Instead, the macro approach focuses on complementary aspects of market power. As such we see the two approaches as inherently complementary; and therefore recommend a joint endeavour in studying sources and implications of market power across industries in the economy.

In particular, the macro approach focuses on economy-wide determinants and implications of market power, in particular those that relate to the distribution of surplus, to output and to welfare. Key in this approach are the effects through prices, even in markets that are not subject to market power. These general equilibrium effects can have a large and lasting impact on households and welfare. Consider for example the case where market power in the goods markets rises substantially. Then even if the labor market is competitive, there are general equilibrium effects on wages. The goods market power leads to higher prices, and hence lower quantities as we move up the demand curve. Lower quantities in turn lead to lower production, and hence a decline in the demand for labor, and thus falling wages for any labor supply that is not perfectly elastic even if the labor market

[1] We will make the point several times that it is the set of facts jointly, of various measures each with their challenges, that paint a robust pattern across a wide range of industries. Of course, this does not imply that all sectors and markets behaved in the same way. In fact a recent literature has started to document industry-specific trends using different approaches, and have found that for instance that markups increased, in the US, in the packaged consumer good industries, beer, cement, while remaining stable in the automobile industry, and even declined in the steel industry.

is perfectly competitive.

These general equilibrium (GE) effects on prices outside the market with market power now have welfare implications that reach beyond the consumer. Taking into account these GE effects and in a micro-founded framework with household preferences, we can now calculate welfare. This includes the effect of the consumer surplus, but now also the welfare effects on all other markets, including the labor market. In fact, there is no such thing as a *consumer welfare standard*, which is often cited in policy circles. There is consumer surplus and total welfare, but no consumer welfare, because welfare takes into account the effect on utility of all agents in the economy, not just the consumers.

In a micro setting of a specific market, the partial equilibrium approach is justified and focuses on the local effect of market power, taking price effects in other markets as given. But this misses out on important implications in a macro setting when market power changes economy-wide, not just in a local market.

In this review, we focus on two key aspects of this macro approach. First, the measurement of market power in the aggregate and the components that contribute to the observed rise in market power since 1980 as well as the role of heterogeneity in aggregation. In particular, we adopt the production approach to measure markups, a leading indicator of market power, which consist of measuring the (inverse) of a variable input's expenditure share and its associated output elasticity.[2] The main insight here is that the output elasticity contributes relatively little to the rise of market power.[3] Most of the rise is due to reallocation – the change in the distribution of shares towards high markup firms – and the expenditure share of variable inputs. This is important, because obtaining the output elasticities requires estimating production technologies which requires functional form assumptions, dealing with input endogeneity and the presence of product differentiation, all while allowing for firms to have market power. We further find that the evolution of the distribution of estimated output elasticities is remarkably similar to that of cost shares. We implement cost shares at the firm level, while these require strong assumptions on the underlying technology and cost function (e.g., all inputs are variable, including capital, as well as constant returns to scale in production), they allow for substantial heterogeneity in technology across firms and time.

Second, we focus on the causes and consequences of the rise in market power. The major determinants of market power are technology (total factor productivity (TFP) and returns to scale), preferences (demand), market structure (the number of competitors and how firms determine prices and quantities) and conduct (how firms compete and whether, for example, there is collusion or common ownership). Research finds that market power is mostly driven by technological change, as well as changes in preferences and conduct. We frame these mechanisms in the context of a model that is standard in the literature and which contains each of these determinants affecting market power. This allows us to investigate how the determinants of market power pass through to prices and markups, output, the labor share and welfare. One of the main takeaways is that market power affects the distribution of outcomes across agents: between wages (the labor share) and profits; between profits of

[2]It is understood that high margins reflect the underlying technology and scale of production, the potential presence of innovation and therefore this indicator needs to be integrated with additional measures of profitability, and are often aided by a specific model to connect it to market power. See e.g. De Loecker et al. (2021).

[3]This is conditional on addressing the set of challenges that go into measuring these elasticities.

dominant firms and startups or new entrants; between wages of production workers and superstar workers. We show how the determinants of the technology transmit to these distributional outcomes, and have economy-wide consequences.

We further use the model to evaluate what aggregate markups can tell us about welfare. In particular, we focus on settings where the aggregate cost-weighted markup, a measure commonly used in macroeconomics, is held constant, while we investigate the impact of changes in any of the other parameters change. Our main insight is that an aggregate markup alone cannot inform us about the state of the economy, whether it be welfare, output, the labor share or any other macroeconomic outcome. All determinants – preferences, conduct, market structure, technology – affect the behavior of firms, and as a result also the underlying outcomes such as markups, output, wages and welfare. The distinguishing feature is the joint distribution across firms of these outcomes. As the distribution of outputs, costs, markups,... changes, so does welfare. Simply focusing on an aggregate cost-weighted markup for example, without further knowledge of the underlying distribution of outcomes, does not inform us whatsoever about welfare. It is important therefore to study the evolution of the underlying distributions to understand what happens to the evolution of market power. Observing a rise in the aggregate markup when all firms are identical has very different welfare properties than observing the same path of aggregate markups where the distribution of markups becomes more dispersed.

In analyzing the effect of market power on wages, we distinguish between the direct effect of monopsony power and the indirect, general equilibrium effect from output market power on wages. In the text we discuss the difference between the mechanisms and refer to the literature that shows that the indirect general equilibrium mechanism is the dominant force that accounts for three quarters of the wage-productivity gap, and one quarter is due to monopsony power.

We finish this review article with a discussion of the policy options to redress the negative effects of market power and improve welfare. The distribution of the surplus between firms and workers plays a key role. The extent to which there are inefficiencies due to market power, necessarily implies that competition policy also has distributional implications. Even though redistribution is often not the mandate or target of competition agencies, at the macro level competition policy has distributional consequences. We discuss conventional policies such as taxation, minimum wages and maximum wage caps for executives, as well as competition policy that has a much broader impact than just on consumer surplus. This opens the discussion whether competition policy can be seen as yet another social policy that achieves distributional goals in the labor market and in the market for startups for examples. We discuss the role of ex ante regulation such as interoperability rather than ex post enforcement to achieve those distributional goals of competition policy.

Related Literature This review article builds on a large body of earlier work. Even though the academic discussion on the evidence on the rise of economy-wide market power is relatively recent, there is by now a substantial literature documenting the facts as well as on the macroeconomic aspects of market power. In addition, there is a long tradition in the industrial organization literature dealing with closely related topics. Below, throughout the text, we refer to that work when we discuss the issues to which they relate. We do not attempt to cover all relevant contributions, instead we focus on

a few key aspects, and how they inform the broader discussion around the role and impact of market power across the economy.

2 Economy-wide Market Power: Facts and Measurement

The main insights of the research on market power in recent years is that market power is pervasive economy-wide, and that there has been a marked increase in various indicators linked to market power, chief among them markups and proft rates, since around 1980. To measure market power, we use individual firm-level indicators that inform us about the potential of pricing power of each firm and its relation to all other firms in the economy. There are three aspects about this evolution that are worth noting. First, the rise in market power is robust when using different methods to measure market power: whether using markups, profit rates or market valuations. Moreover, different estimation methods for markups yield similar conclusions, for example using the demand versus the cost approach; or using different techniques to estimate output elasticities. Second, market power – and the increase over time – is all about the change in the *distribution* of these individual indicators. In particular, we tend to observe a marked increase in the upper percentiles of the distribution of the chosen indicator (markup, profit rate, firm value), while the median tends to remain constant and lower percentiles tend to decline. In other words, the rise of market power is highly beneficial to a small number of dominant firms, and detrimental to the majority of the remaining firms. As mentioned above, we do imply that this is true of literally all sectors of the economy, in fact the cross-industry heterogeneity is a natural feature of relying on micro data as we do in our analyses, and we see the recent arrival of industry-specific analysis of market power over time as a highly valuable, and complementary effort.[4]

2.1 Measuring Market Power in the Macroeconomy

Market power is complex. It is simultaneously determined by different forces: demand and consumer preferences, technology and costs, market structure (such as the number of competitors and the potential for entry) and conduct (the degree to which firms are able to collude for example, or the extent to which there is common ownership of competitors). We discuss four commonly used measures of market power: concentration index, markups, profit rates, and market value.

Concentration Index. Concentration indices are summary statistics of the shares in a defined market. Those market shares can be of revenue, of employment, or of any measure that reflects the relative size of firms. The most used concentration measure is the Herfindahl-Hirschman Index (HHI) which is equal to the sum of the squared market shares.[5] Key is that the HHI varies simultaneously with the number of competitors – all else equal, fewer competitors implies a higher HHI – as well as

[4]For example, rising margins have been documented in the US brewing and cement industries, and in the overal market of consumer products (De Loecker and Scott (2025), Miller et al. (2025), Brand (2021), Döpper et al. (2025), while margins were found to be relatively stable in the US automobile industry (Grieco et al. (2024)), and declining in the US steel industry (Collard-Wexler and De Loecker (2015).

[5]For market shares expressed in percentages, HHI ranges from 0 under perfect competition to 10,000 under monopoly.

with the distribution of market shares – more inequality also implies a higher HHI – which means that the HHI is akin to summary measures of inequality such as the variance, or the Gini coefficient. In a similar vain, concentration ratios ($C(n)$) measure the share of the largest n firms in a market, in other words, the upper tail of the size distribution.

Concentration measures, especially HHI, are extremely popular in policy because they are easy to understand and can collapse the complexity of market power in a simple statistic. Judges tend to apply merger guidelines where HHI larger than 3000 flags a concern of excessive market power. Despite their popularity, there are however serious concerns regarding the reliability of concentration measures as indicators of market power, most vocally expressed by Bresnahan (1989). First, it requires a definition of what the market is and which firms are included as competitors. There are many causes for concern. Excluding/including even only one more firm drastically changes the HHI. Not surprisingly, in a merger case, merging firms argue that many firms are competitors while the competition authority argues the opposite. But this is beyond the point. There is no binary rule whether two firms are in or out of the same market. Rather markets are continuous and depend crucially on the degree of substitutability of products. The less substitutable goods are, the more market power firms (potentially) can execute. Conceptually, all firms are in the economy, and are in the market, and there is continuum of degrees of interaction that is more or less relevant for market power they exert. That is why academic research has focussed on measures of market power that derive from more structural interpretations of the interaction between firms.

In Figures B.1 and B.2 in the Online Appendix, we plot the HHI and C(4) for the Compustat data.[6] In both there is a sharp decline from the beginning of the sample until the mid 1990s, followed by an increase in both concentration measures. The rise of concentration in the last three decades is in line with the findings by Grullon, Larkin, and Michaely (2019) and Barkai (2020), though the sharp decline during the 1980s and first half of the 1990s may at first sight contradict the rise in market power that we find from other measures (markups, profits, market value). But this also highlights the limitations of concentration measures that depend on the number of firms in each sector. In the same figures we plot the number of firms, and we see a sharp increase in the number of firms, which mechanically leads to an decline in HHI.[7]

Markups. Economists also use the notion of markup, the ratio of the price to the marginal cost, as a measure of market power. Markups are a measure of the pricing power of firms, as it naturally follows from the firm's optimization decision that equates marginal revenue with marginal cost. There are different approaches to obtain the markup, and they can be categorized in two classes: the demand-conduct approach (see among many others Bresnahan (1989) and Berry, Levinsohn, and Pakes (1995)) and the production approach, initially proposed by Hall (1988) using aggregate data and taken to micro data by De Loecker and Warzynski (2012). Refer to De Loecker (2011) for an early discussion on the various approaches, and to De Loecker and Scott (2025) for a recent discussion and implementation of two structural approaches that both take the starting point that marginal costs are

[6] Keeping in mind all the challenges that go into measuring HHI indices in the sample of listed firms.

[7] The rising number of firms has also led opposing evolutions of concentration measures at different levels of aggregation, see Rossi-Hansberg, Sarte, and Trachter (2021) for example and Eeckhout (2020) who points out the mechanical relation driven by the number of firms.

not observed. This is in contrast to the accounting approach where marginal costs is equated to average costs, and therefore markups are directly read from the data by considering the ratio of revenue to total (variable) costs.

Following the production approach of De Loecker and Warzynski (2012), from the first-order condition of cost minimization (with respect to a variable input of production V), that we can write the markup, for firm i at time t as:

$$\mu_{it} = \theta_{it}^{V} \frac{R_{it}}{P_{it}^{V} V_{it}}. \tag{1}$$

where θ_{it}^{V} is the associated output elasticity of that variable input V, and $P_{it}^{V} V_{it}$ its observed expenditure, while R_{it} is observed revenue.

Following De Loecker, Eeckhout, and Unger (2020), we plot the baseline measure of markups for the US economy based on the publicly traded firms in Figure 1, which illustrates that the rise of market power starting in 1980 is driven by the increase in the dispersion, and in particular by the increase in the top percentiles of the markup distribution.

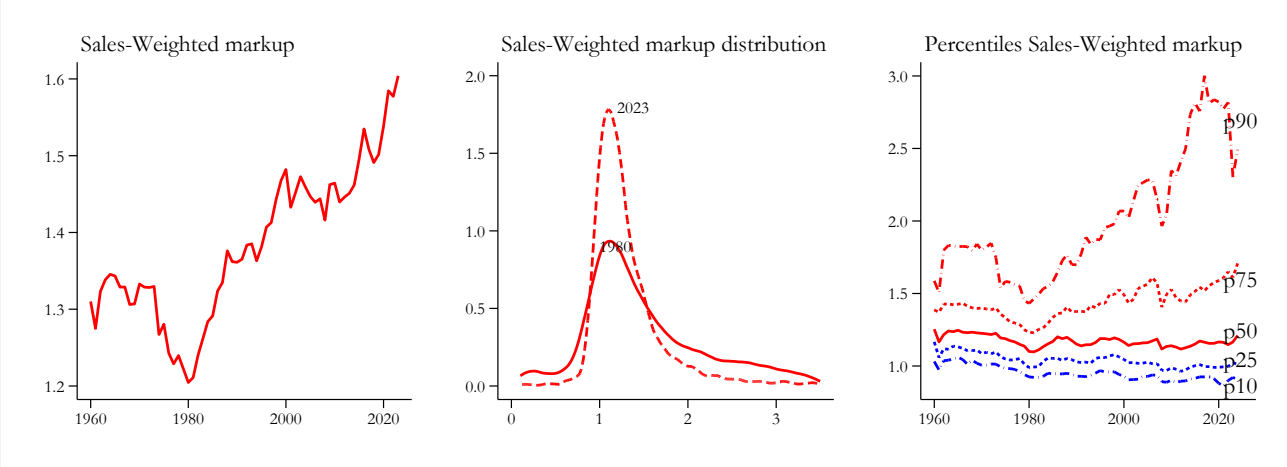


Figure 1: Sales-weighted markup over time and distribution

Note: The center panel limits the range of the plot of the markup within [-0.2, 0.3]. In the right panel, blue indicates below-median percentiles and red indicates above-median percentiles.

Below, we elaborate on issues of measurement and estimation of markups.

Profit rates. While the markups measure the ability of the firm to price its output above the marginal cost of production, they do not include any overhead cost such as sales and marketing, investment (measured by the user cost via depreciation and amortization), financing cost of debt or taxes. Instead, profits take into account those overhead costs. A firm with overhead costs will have markups larger than one even if it is operating perfectly competitively and has zero profits. And it may increase its markups when overhead costs rise in order to avoid making losses.[8]

[8]The role of overhead is often measured in accounting data as SG&A, Selling, General and Administrative Expenses. De Loecker, Eeckhout, and Unger (2020) document a rise in overhead during the 1980s and 1990s. Olmstead-Rumsey (2019); De Ridder (2024) show that intangibles play an importnat. See also Aghion, Bergeaud, Boppart, Klenow, and Li

The economic profit rate, measured as the ratio of economic profits as a share of revenue, is therefore an additional measure of market power that informs us about the flow of income a firm generates, including overhead.[9] We use two measures of economic profit rates.

Profits are equal to Revenue minus Total Cost (TC_{it}). For the two definitions of economic profits, we use the accounting definition TC_{it}^{acc} and the imputed definition TC_{it}^{imp}. The accounting profit and total cost definition is (see also Eeckhout (2025)):

$$TC_{it}^{acc} = COGS_{it} + SGA_{it} + r_{it}TA_{it} + DEPAM_{it} - NOPI_{it} - SPI_{it} + TAX_{it} \tag{2}$$

where $COGS_{it}$ is cost of goods sold, SGA_{it} is Sales, General and Administrative Expenses, $r_{it}TA_{it}$ is the interest rate times total assets,[10] $DEPAM_{it}$ is depreciation and amortization, $NOPI$ is non-operating income/expenses, SPI is special items,[11] and TAX_{it} is all tax liabilities. All nominal variable are deflated using the GDP deflator, as above. Because the user cost in accounting terms does have a price for equity, we impute the user cost of the firm's equity by using the average rate that the firms pays on all its debt.[12]

Alternatively, we define total costs by imputing the user cost of capital using the value of the stock of physical capital as well as the risk-free interest rate, inflation and imputing a cost of depreciation. We use this definition in our previous work, De Loecker et al. (2020):

$$TC_{it}^{imp} = COGS_{it} + SGA_{it} + (0.12 + i + r^f) \times PPEG_{it} \tag{3}$$

where PPEG is the accounting measure for physical capital, i is the inflation rate (measured as the GDP deflator), and r^f is the risk-free rate, measured as the 10-year treasury yield and the imputed depreciation/amortization is 12%.

Economic profits based on accounting data (see Eeckhout (2025)), are plotted in Figure B.3 in the Appendix, and economic profits based on and imputed user cost of capital (see De Loecker, Eeckhout, and Unger (2020) and the references therein), in Figure B.4. Both average profit rate increase over the sample period between 1980 and today. The accounting profit rate increases from 2% in 1980 to over 6% now, and the profit rate with imputed user cost of capital starts at 4% in 1980 to over 9% now. The main evolution is the increase in the dispersion, with rising upper and lower percentiles, including more firms with negative profits.

(2023) and Eeckhout and Veldkamp (2022). Traina (2018) uses the total cost, including overhead and capital to calculate markups, which in that case are equivalent to profit rates. For a discussion and evidence on the relation between the markup and the profit rate, see De Loecker, Eeckhout, and Unger (2020), Syverson (2019) and Basu (2019).

[9] Of course, it inherits all the challenges of relying on reported cost items. This therefore constitutes a clear trade-off between not having any measure of profits, in the absence of imposing a full model of entry and exit, to recover measures of economic fixed costs.

[10] Observe that $rTA = r(Debt + Bookvalue) = INT + rBookvalue$ were INT are the interest payments on debt.

[11] Non-operating income results from secondary business-related activities, excluding those considered part of the normal operations of the business, like dividend income, rental income, foreign exchange adjustment,... Special items are unusual or nonrecurring items presented before taxes by the company. For example, flood, fire, and other natural disaster losses; impairment of goodwill, unamortized intangibles,...

[12] This total cost differs from the accounting total cost since accounting costs do not value the opportunity cost of equity kept in the firm. To that effect, instead of adding the total interest payments INT_{it}, we add the user cost of total assets, $r_{it}TA_{it} = r_{it}Bookvalue_{it} + INT_{it}$ since $TA_{it} = Bookvalue_{it} + Debt_{it}$ and where $r_{it} = \frac{Debt_{it}}{INT_{it}}$ is the total debt divided by the interest payments.

Market values. Finally, while profits include a measure of overhead and costs of investment, they are still a flow of income the firm generates. But profits may be delayed. The very nature of investment is to generate future profits. Of course, suitably adjusting the user cost of capital (with depreciation and amortization), accounting profits aim to attribute the investment proportionally to the period in which those investments are used, this is an imperfect measure. For publicly traded firms, the stock market value represents the market's valuation of the future flow of all dividends and therefore all future profits and equity embedded in the firm.

In Figure B.5 in the Appendix, we plot the average market value, which increases tenfold between 1980 and now. As with markups and the profit rate, this is mainly driven by a sharp increase in the dispersion. The distribution is approximately lognormal and shifts to the right with increasing variance, even in log valuations. The upper percentiles increase markedly, most notably the 90th percentile. The lower percentiles remain fairly stable. This confirms that the rise of market power in the aggregate is really about a change in the distribution of dominant positions. A few firms achieve dominant positions with associated market valuations, while most firms see little increase or even a decrease in their market valuation. This is also evident from observing the evolution of the market valuation of the top 7 publicly traded firms (see Eeckhout (2025)): in 2023 those 7 firms account for around one quarter of the total market capitalization of *all* publicly traded firms, and they are growing even faster as they account for just under half of the growth of the total stock market in 2023.

2.2 The Components of Markups

To calculate the markups following the production approach, we need to make assumptions about firm behavior, the production technology and a notion how the variable factor market clears. Relying on firms to minimize cost when selecting at least one variable input, over a given time frame which is typically one year in the data, we can estimate markups by relying on observed expenditures on the variable input, revenue and estimates of the relevant factor's output elasticity.

The markup in equation (1) consists of the product of two components, the output elasticity of the variable in put θ_{it}^V and the inverse expenditure share of that variable input $\frac{R_{it}}{P_{it}^V V_{it}}$. We obtain these two components as follows. We need accounting or survey data on the revenue and on the expenditure on the inputs of the firm, including on the variable input. With cost data, we can immediately construct the expenditure share. We also need the output elasticity, which requires assumption on the production technology. Directly or indirectly, that involves estimating a production function.

There are plenty of intricacies regarding the estimation of production functions, necessary obtain output elasticities. [13]

Because we are interested in the evolution over time, we want to allow for time-varying production functions, that is with time-varying parameters. That is, we estimate a production technology

[13] See among many others De Loecker and Scott (2024), Ackerberg and De Loecker (2024) and De Loecker and Syverson (2021) for an overview. For individual studies see Klette and Griliches (1996); Olley and Pakes (1996); De Loecker and Warzynski (2012); De Loecker, Eeckhout, and Unger (2020); Gandhi, Navarro, and Rivers (2020); Ackerberg, Caves, and Frazer (2015); Ackerberg, Benkard, Berry, and Pakes (2007); Bond, Hashemi, Kaplan, and Zoch (2021); De Ridder, Grassi, and Morzenti (2025, forthcoming); Hubmer, Chan, Ozkan, Salgado, and Hong (2024)

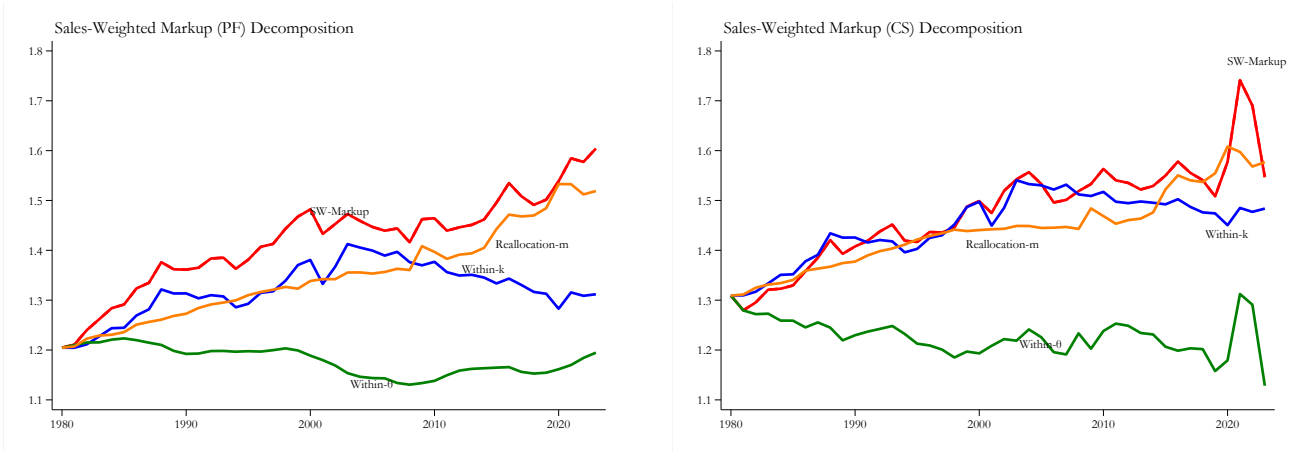
for every year (or for a moving window of time around every year).[14]

Figure B.6 in the Appendix plots the evolution of the aggregate output elasticities estimated using different methods, for two cases, one where the aggregate elasticities are unweighted (left panel) and one where they are weighted by the sales share of the firms in the sample (right panel). In both cases, there is a plot of the benchmark estimated output elasticity which is estimated for a five-year rolling time window, and the aggregate cost share, which is is the variable input's expenditure share of all costs. Both elasticities are fairly stable, with a minor decline by no more than 5 percentage points.

The Contribution to Markups of Expenditure Shares versus Output Elasticities. To learn what drives the rise in markups, we decompose the growth of markups into its different contributing components. Building on Haltiwanger (1997) we follow De Loecker, Eeckhout, and Unger (2020) to decompose the revenue-weighted markup:

$$\mu_t = \sum_i m_{it}\,\theta_{it}\kappa_{it} \tag{4}$$

where $m_{it} = \frac{R_{it}}{\sum_{i\in N_t} R_{it}}$ is the revenue share and $\kappa_{it} = \left(\frac{R_{it}}{P^V_{it}V_{it}}\right)$ is the inverse expenditure share on the variable input. In Appendix C we derive in detail the decomposition in terms of the change due to change in the revenue share Δm, due to the change in the output elasticity $\Delta\theta$, and the change due to the inverse expenditure share $\Delta\kappa$.



(a) Production Approach

(b) Cost Share

Figure 2: Decomposition of the Change in Sales-Weighted Markups - Production Approach vs Cost Share

Note: The figure does not plot the second- and third-order interaction terms, such as Cross(m,θ), Cross(m,κ), Cross(θ,κ), and the triple interaction term. Those terms are quantitatively negligible. The net entry terms are included in each of the components.

Figure 2a plots the cumulative evolution of each of the components that contribute to the change in the revenue-weighted markup. The reallocation component (Δm) is by far the largest and consist of

[14] Of course, production technologies may be time-varying even if the parameters are constant, simply because due the inputs are time-varying, which is the case for example for the translog production technology.

a significant component of entering firms that capture higher market shares. The within component due to the change in the inverse revenue share of the variable input ($\Delta\kappa$) accounts for about one third of the rise in markups. Finally, the within component due to the change in the output elasticity ($\Delta\theta$) does not contribute at all to the rise in aggregate markups. If anything, the contribution of the output elasticities is negative. This is robust to whether the output elasticities are obtained using production function estimation or using cost shares (Figure 2b).

This establishes that the main finding that the rise in the aggregate markups is completely driven by the rise in the reallocation term as well as the inverse expenditure share on the variable input, and not by the change in the output elasticity. Different output elasticities may show up in different levels, but they do not show a sharp rise that can account for the rise in markups.

Estimating Output Elasticities: Production Function versus Cost Shares. In addition to the variation in the output elasticity over time, there is variation in the cross section across industries and firms. A production function estimation typically exploits the variation across firms to pin down the parameter values of the technology, thus obtaining production function parameters that are common across firms (across some pre-assigned level of aggregation). Due to the nature of pooling across firms deemed to produce under a common production technology, while allowing their efficiency to vary freely across firms, this approach may potentially leave heterogeneity in technology unaccounted for. This may then result in measured markup heterogeneity within industries.

To address this concern, some approaches generate firm-specific technology estimates, that is, each firm has its particular output elasticity. This is either obtained through a non-parametric production function or at least a flexible functional form that allows for sufficient variation, such as the translog production function considered in De Loecker and Eeckhout (2017). An alternative to obtain firm-specific elasticities, and easy to implement, is the factor-specific cost share, measured as the ratio of the variable input's expenditure over the cost of the inputs in production[15] (see amongst others Syverson (2004); Foster, Haltiwanger, and Syverson (2008)), or it has a particular structure where the output elasticity is derived from the firm-level first order condition (see Deb, Eeckhout, Patel, and Warren (2024, 2022); Hubmer, Chan, Ozkan, Salgado, and Hong (2024)). Using the cost share as the output elasticity is particularly transparent and simple. That does not mean that there are no underlying assumptions and structure. To the contrary, we know that the output elasticity is equal to the cost share only if all inputs, including capital, are variable, and if there are constant returns to scale. In addition, cost of production is readily available in the data.[16]

Now with output elasticities derived from cost-shares, we obtain variation in markups that is potentially different from the variation in markups due to output elasticities that are common to all firms in one industry. Still, when we plot the markup distribution comparing markups with

[15] We use COGS and our measure for the user cost of capital as our measure of the cost of the inputs of production, but not overhead since overhead is not an input in production and therefore does not affect the output elasticity. If we use all costs, including overhead, then the markup is closely related to the profit rate as a share of sales: $\pi_{it} = 1 - \frac{1}{\mu_{it}}$. Profit rates are of different magnitude than markups, and have increased equally sharply from 2% in 1980 to 8% now (see De Loecker et al. (2020)). Foster et al. (2024) use establishment data from the US Census of Manufacturers and calculate markups using cost shares based on total costs.

[16] Syverson (2004), and De Loecker and Syverson (2021), and Hubmer, Chan, Ozkan, Salgado, and Hong (2024) proposes a method that allows for increasing returns to scale.

estimated output elasticities versus those derived from cost shares, there is no substantial difference in the distribution. See Figure B.7 in the Appendix. In the case of cost shares, the distribution is slightly shifted to the right in all years reflecting the slightly higher level of the markup under cost shares. With cost shares, elasticities are slightly higher (see Figure B.6 in the Appendix) and as a result the markup is higher. Most importantly, the shape of the distribution is remarkably similar. This is also evident from the evolution of the percentiles of these distributions, which show a similar pattern.

We repeat the same exercise for the accounting measure of costs (Figures B.9 and B.8 in the Appendix), and find that the average markup has a constant level difference, but shows an otherwise similar evolution. The distributions evolve similarly over time as is evident from the evolution of the percentiles.

Despite the fact that the cost shares measure a firm-specific output elasticity, and the output elasticities obtained from an estimated production function is common to all firms in the industry, those two elasticities and therefore the markups are highly correlated. Figure B.10 in the Appendix shows the scatter plot of both markups for 2019, one obtained with the estimated output elasticities and one obtained with the cost shares. The markups line up with an R-squared of 0.84. This is remarkable because the output elasticities are common for all firms in an industry, whereas cost shares are firm-specific. To there is a lot more variation in the elasticities based on cost shares than there is variation in the output elasticities. Nonetheless, the markups are similar. This confirms that most of the variation in markups is due to the variation in the expenditure share of the variable input, and not due to the variation in the output elasticity.

Moreover, there is not a particularly systematic pattern in the relation between the markup based on cost share elasticities and estimated elasticities. Figure B.11 in the Appendix shows that there is a slightly negative relation between the ratio of the two markups and the level of the markup, and there is no systematic relation between the ratio of the two markups and revenue.

2.3 Aggregate Markups and Heterogeneity

The four main determinants of market power – preferences, technology, market structure and conduct – jointly pin down the distribution of market power in the economy. This has implications not only for the distribution of markups, but also for the distribution of firm size, the distribution of prices, and welfare.

Those four determinants, through the endogenous distributions of outcomes, then also affect aggregate markups, depending on how they are weighted. Two commonly used aggregates are revenue-weighted (or sales-weighted) and cost-weighted markups. In the macroeconomics literature, the cost-weighted markup is often preferred due to its properties for aggregation, see for example Grassi (2017); Edmond, Midrigan, and Xu (2023).

Aggregate Markups. First, we define the different aggregate markups. The Revenue-weighted aggregate markup is:

$$\mu_t^R = \sum_i s_{it}\mu_{it} \tag{5}$$

where $s_{it} = \frac{R_{it}}{\sum_i R_{it}}$. The Cost-weighted aggregate markup:

$$\mu_t^C = \sum_i c_{it}\mu_{it} \tag{6}$$

where $c_{it} = \frac{p_{it}^V V_{it}}{\sum_i p_{it}^V V_{it}}$.[17] Finally, the Revenue-weighted harmonic mean aggregate markup:

$$\mu_t^{RH} = \left(\sum_i s_{it}\mu_{it}^{-1}\right)^{-1} \tag{7}$$

where $s_{it} = \frac{R_{it}}{\sum_i R_{it}}$

The markup, as defined in our measurement section above, is the ratio of the firm's price to its marginal cost. Under constant returns to scale, we can also write the markup as the ratio of revenue to total cost[18]:

$$\mu_{it} = \frac{R_{it}}{TC_{it}}. \tag{8}$$

To see that, observe that under constant returns marginal cost is equal to average cost: $\lambda_{it} \equiv \frac{dTC_t}{dy_t} = \frac{TC_{it}}{y_{it}}$ (and thus $TC_{it} = \lambda_{it} y_{it}$). Then since markup is the ratio of price to marginal cost, it follows that under constant returns and no scale economies, $\mu_{it} = \frac{p_{it}}{\lambda_{it}} = \frac{p_{it}y_{it}}{\lambda_{it}y_{it}} = \frac{R_{it}}{TC_{it}}$. Note that it is important to distinguish between constant returns – that doubling the inputs of production doubles output – and scale economies – that there are no fixed costs or overhead – as both need to be satisfied to ensure that the marginal product is equal to the average product. Equivalently then, since $R_{it} = \pi_{it} - TC_{it}$, we can write profits as $\pi_{it} = R_{it} - TC_{it} = R_{it} - \lambda_{it}\frac{R_{it}}{p_{it}} = R_{it}\left(1 - \frac{1}{\mu_{it}}\right)$, which implies (8).

For aggregation purposes and to be able to express aggregate markups, costs, revenue,... in a similar way as those firm-level variables, in the macro literature we often encounter the following definition of the aggregate markup μ_t^{macro}

$$\mu_t^{Macro} \equiv \frac{R_t}{TC_t} = \frac{\sum_i R_{it}}{\sum_i TC_{it}} = \sum_i \frac{p_{it}y_{it}}{\sum_i TC_{it}} = \sum_i \frac{\mu_{it}\lambda_{it}y_{it}}{TC_{it}}\frac{TC_{it}}{\sum_i TC_{it}} = \sum_i \mu_{it}\frac{TC_{it}}{\sum_i TC_{it}} \tag{9}$$

which is the total cost-weighted markup under constant returns to scale. We can further write this macro markup as:

$$\mu_t^{Macro} = \frac{\sum_i R_{it}}{\sum_i TC_{it}} = \frac{\sum_i R_{it}}{\sum_i \lambda_{it}y_{it}} = \sum_i \frac{R_{it}}{\sum_i \mu_{it}^{-1}p_{it}y_{it}} = \left(\sum_i \mu_{it}^{-1}s_{it}\right)^{-1} = \mu_t^{RH}. \tag{10}$$

The macro markup is equal to both the Cost-weighted markup and the Revenue-weighted harmonic

[17] The cost weight can also be based on broader costs that include other variable inputs, capital, overhead,... beyond the variable input V_{it} that is used to obtain the markup.

[18] All inputs are required to be variable inputs in production.

mean markup. Key in this derivation is the assumption of a constant returns to scale technology and the absence of scale economies. Without, the cost-weight of the variable input and of the total cost here are not equivalent. Moreover, then the output elasticity does not aggregate.

To see this, observe that the cost-weighted aggregate markup μ_t^C can be written as:

$$\mu_t^C = \sum_i c_i \mu_i = \sum_i \frac{P_{it}^V V_{it}}{\sum_i P_{it}^V V_{it}} \theta_t^V \frac{R_{it}}{P_{it}^V V_{it}} = \frac{\sum_i P_{it}^V V_{it}}{\sum_i TC_{it}} \frac{\sum_i R_{it}}{\sum_i P_{it}^V V_{it}} = \frac{R_t}{TC_t} = \mu_t^{Macro}. \tag{11}$$

The latter equation only holds under constant returns to scale, and when the output elasticity is common to all firms ($\theta_{it} = \theta_t$) and equal to the aggregate cost share.[19]

Under the assumptions of constant returns to scale, the distribution of firms (productivities, revenue shares, cost shares, markups,...) still matters for the aggregate markup. The dispersion of the cost share distribution or of the revenue distribution affects these weighted aggregate markups. In fact, Grassi (2017), Proposition 4, links the aggregate markup to the concentration (HHI or Hannah and Kay (1977) concentration index).

Aggregate Markups and Heterogeneity. Most importantly, any aggregate markup – whether it is the Cost-weighted, the Revenue-weighted or the Revenue-weighted harmonic mean aggregate markup – cannot be taken as a summary statistic that is independent of the underlying distributional characteristics. However appealing any aggregate markup is, it can never tell the full story. In other words, for any aggregate markup, there are a continuum of economies that generate that same markup but that each have different output, wages and welfare properties. This is nothing different from stating that there are a continuum of distributions that generate the same variance, but that have different means. In other words, all moments matter to fully describe the distribution, and in our case at hand, the entire distribution of markups matters for describing the phenomenon of market power. We can therefore not sustain that an economy with a certain aggregate markup has certain welfare properties. Without further information, we simply don't know. Two economies with the same cost-weighted markup for example can have very different welfare properties. The same holds true for two economies with the same revenue-weighted markup; or any aggregate markup for that matter.

This is already transparent from immediately observing the expression for the cost-weighted aggregate markup:

$$\mu_t^C = \sum_i c_{it} \mu_{it} = \sum_i \frac{p_{it}^V V_{it}}{\sum_i p_{it}^V V_{it}} \mu_{it} = \sum_i \frac{p_{it}^V V_{it}}{\sum_i p_{it}^V V_{it}} \theta_t^V \frac{R_{it}}{p_{it}^V V_{it}} = \theta_t^V \frac{\sum_i R_{it}}{\sum_i p_{it}^V V_{it}}. \tag{12}$$

It is equal to the average revenue over the average expenditure on the variable input in the economy,

[19] Observe that if the output elasticity θ_{it}^V is firm-specific (as is the case with cost shares), then the cost-weighted markup is equal to

$$\mu_t^C = \sum_i c_i \mu_i = \sum_i \frac{P_{it}^V V_{it}}{\sum_i P_{it}^V V_{it}} \theta_{it}^V \frac{R_{it}}{P_{it}^V V_{it}} = \frac{\sum_i \frac{P_{it}^V V_{it}}{TC_{it}} R_{it}}{\sum_i P_{it}^V V_{it}} \neq \frac{R_t}{TC_t} = \mu_t^{Macro}$$

multiplied by the common output elasticity. If all firms are identical and have the same revenue and expenditures on variable inputs, then that yields a cost-weighted markup that is the same as an economy with dispersed revenue and dispersed expenditures. In other words, the cost-weighted aggregate markup does not capture any notion of dispersion of markups in the economy. And the dispersion is exactly what has increased. After all, market power is typically driven by a few dominant firms, and that is precisely what we observe in the evolution of the markup distribution, where the upper percentiles increase since 1980.

The weights matter. Firms that have market power have higher markups, and therefore they tend to have higher revenue relative to costs. As a result, in the distribution they have higher revenue shares and than cost shares. This drives a wedge between the revenue-weighted aggregate markup and the cost-weighted aggregate markup, where revenue gives a higher weight to high markup firms than costs do. The gap between revenue-weighted and cost-weighted aggregate markups has grown from nearly no gap (around 1.20) in 1980 to a gap of 20 points (1.60 versus 1.40). See De Loecker, Eeckhout, and Unger (2020) and Hall (2018). This is informative about what happens to the joint distribution of markups, revenue and costs. In particular, it tells us that the distribution of revenue has become much more dispersed than the distribution of costs. High markup firms tend to have disproportionately bigger revenue weights than cost weights, indicating more market power. Because the cost-weighted aggregate markup is independent of the distribution of costs (and of revenue), see equation (12), the cost-weighted markup captures the level effects of higher markups across the board, but not how those are distributed.

The relative difference between the revenue-weighted and cost-weighted markup measures this gap. Edmond, Midrigan, and Xu (2023) points out that this gap, under some assumptions, most notably a common output elasticity, is equal to the variance of the markup:

$$\frac{\mu^R - \mu^C}{\mu^C} = \text{Var}_c \left[\frac{\mu_i}{\mu^C} \right]. \tag{13}$$

The heterogeneity thus has implications for the welfare impact of markups. Even if the cost-weighted aggregate markup stays constant, welfare varies with the distribution of markups. To evaluate welfare, it is important to have additional information on the distribution, and on the determinants of market power that drive the distribution of markups: preferences, technology dispersion, returns to scale, market structure and conduct.

Below, we evaluate the implications of heterogeneity and how aggregate markups inform us about the underlying economy in the context of a model.

3 The Macroeconomic Causes and Consequences of Market Power

The economy-wide presence of market power has macroeconomic implications. The mechanism is through prices, such as wages, input prices, output prices,..., and which in turn affects decisions such as demand, labor supply, entry, investment,... In other words, market power that changes economy-wide has general equilibrium effects. When focusing on specific, well-defined markets, and given the

difficulties analyzing strategic interaction between firms in large economies, researchers often ignore those general equilibrium effects or assume them away. This helps them make progress understanding the details and intricacies of markets and individual firms, and their is great value in pursuing that strategy in some instances, and for certain specific research questions.

In the light of the pervasive evidence that measures of market power is evolving economy-wide, instead our approach is to analyze the macroeconomics of market power through the general equilibrium effects it has economy-wide. With the evidence from the rise of variable profits margins and profits, that we have discussed in the former section, we now describe the mechanisms through which market power has macroeconomic effects. The main objective is to evaluate the implications of the sources of market power on welfare. First, we discuss the main causes and consequences, then we present a model that formalizes these mechanisms and analyze the comparative statics effects on output, wages, profits and welfare. We then discuss the relation between welfare and the properties of the economy when aggregate markups are kept constant. The objective is to highlight that aggregate markups alone do not inform us about welfare or the aggregate economy in general. We need more detailed distributional information.

3.1 Causes and Consequences

Causes. Economy-wide market power is affected by different forces, the most important are technology, demand, market structure and conduct. Technology is probably the broadest as it includes the Total Factor Productivity (TFP), returns to scale and fixed costs (scale economies). The heterogeneity of TFP in particular plays a central role as it differentiates the firms and thus affects the distribution of outcomes. Demand factors are often summarized by the demand elasticity which is derived from the preferences that households have over the different goods, governed by the elasticity of substitution between goods. The market structure mainly represents how many firms compete and how. This is of course also affected by technology as investment and entry costs determine how many firms enter the market. Conduct represents the extent to which behavior by firms is affected by their ability to exert market power, including collusion or common ownership which allows firms to determine price setting across competing but commonly owned firms.

There is evidence that each of these causes contribute to the change in market power. De Loecker, Eeckhout, and Mongey (2021) and Edmond, Midrigan, and Xu (2023) find that technology in particular plays a prominent role.[20] This also helps understand why markups make such a sharp increase in 1980. This is at the time when digital technology – the use of desktop computers in households and firms, servers and mainframes for data processing,... – starts to become economically significant. They find that starting in 1980, the technology shows an increase in the dispersion of TFP, as well as an increase in the fixed costs. The increase in the dispersion of technologies is robust to different methods of estimating the technology, see for example Deb, Eeckhout, Patel, and Warren (2022). Recent work by Hubmer, Chan, Ozkan, Salgado, and Hong (2024) finds that there is also variation in returns to scale at the firm level, further contributing to the view that dispersion in technologies

[20] Sutton (1991, 2001) shows that the technological differences are in part due to investment. Firms incur an upfront to invest in increasing their productivity, which allows them to gain market share and increase markups.

across individual firms is a major driver or market power.

Some work documents that there is a change on the demand side. Firms need to build up customer capital, not just produce output. For example, Chiavari (2024) analyzes the role of customer capital in conjunction with returns to scale in building market power. Foster, Haltiwanger, and Syverson (2016) argue that demand and customer acquisition is the main source of size differences between firms, more than technology differences. There is also evidence that preferences are changing over time. In particular, document a decline in substitutability between consumer goods, which results in higher markups (see Döpper, MacKay, Miller, and Stiebale (2025) and Brand (2021)). Casal (2025) provides a rationale for this phenomenon base on investment by firms to make the goods less substitutable. Again, at the sharp increase in 1980, effective market size changed substantially with the intensifying of world trade and ultimately the process of globalization.

Also policy plays a role in the sense that competition policy affects markups because the objective of competition policy is often to redress conduct that is market power increasing. Competition policy is thus one of the determinants of market power.

Consequences. The macroeconomic consequences of market power consist of the same direct effects as well as the indirect effects through the prices system. We first consider the indirect, general equilibrium effects that are typically not analyzed in a micro setting. In the context of the model that we present below, we illustrate some of these effects. These mechanisms are in different forms present in De Loecker, Eeckhout, and Mongey (2021) and we base our discussion on those results.

Wages and the Labor Share. Economy-wide output market power affects wages in general equilibrium. If only one atomless firm sees a rise in output market power, there is no effect on wages.[21] But when market power is economy-wide and a positive measure of firms has output market power, then the rise of market power with rising markups, leads to lower demand and hence fewer quantities of the goods sold. This in turn lowers the units produced and therefore leads to a decline in the demand for labor. As long as the labor supply is not perfectly elastic – meta studies put the labor supply elasticity around 0.25, see Chetty, Guren, Manoli, and Weber (2011) – the lower demand for labor leads to falling wages and falling employment. Quantitatively, the effect on wages is sizable, and contributes to the rising gap between labor productivity and wages.

The decline in wages in conjunction with the decline in employment leads to a declining expenditure on labor (wages times employment). As a result, the labor share – the expenditure on labor as a share of GDP – falls. There is ample evidence of the falling labor share (see amongst other Karabarbounis and Neiman (2014), Barkai (2020), Kehrig and Vincent (2021), Autor, Dorn, Katz, Patterson, and Van Reenen (2020), De Loecker, Eeckhout, and Unger (2020)). This is the case for a wide range of countries (Karabarbounis and Neiman (2014) and Biondi, Inferrera, Mertens, and Miranda (2023)) and coincides with the rise of market power.

There is a direct relation between the increasing profits (or markups) and the decline in the labor

[21] Provided there is no labor market power (see the discussion on monoposony below) and provided there is no granularity such that no one firm is large enough to have an aggregate impact (see Gabaix (2011)) or there are aggregate effects through the input-output network (see Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012)).

share. As market power increase, say the markup increases, then the quantity of the variable input declines. This can be seen immediately from the first-order condition of cost minimization of the firm that leads to the expression for markups in equation (1): if we analyze labor as the variable input, then the labor share is proportionate to the inverse of the markup. This is indeed borne out in the data as De Loecker, Eeckhout, and Unger (2020) show. Even though capital is hardly a variable input, Barkai (2020) shows that also the capital share has declined. This is consistent with the fact that capital and labor are complementary inputs in production. Despite the lumpiness of capital in the adjustment process, as labor declines, eventually so does also capital.

Output and Welfare. In broad terms, GDP consists of capital, labor and profits. As a result of the rise in market power, there is a decline in the labor and capital shares, with a corresponding rise in profits. Indeed, as we have documented above, we also see the rise of profits. At face value, this is a mere redistribution of the surplus. More to the owners of the firms, less to the workers, but since households own the firms, at least in a representative household the fall in labor income is compensated by the rise in profits.

However, there is a net loss, both in output and welfare. The net loss is due to the fact that firms produce less output at higher prices. They get higher revenue, but the deadweight loss due to lower quantities is borne by the households. There is thus not only a redistribution of surplus, there is also a decline in surplus, both of output and of welfare. Estimations of the losses are of similar magnitudes for output and for welfare, and coincide across different estimations. De Loecker, Eeckhout, and Mongey (2021) and Edmond, Midrigan, and Xu (2023) put the output and welfare loss in the range of 8 to 10% per year. That is, GDP and welfare are 8-10% lower today, compared to having the degree of competition that we had in 1980. This loss is large. Just as a comparison, the estimates of the welfare cost of inflation are in the range of 0.5%, around 20 times smaller than the estimated welfare cost of the rise of market power.[22] Economists have given ample importance to the welfare loss of inflation, and it has resulted in an independent central bank to ensure monetary policy and price stability are disconnected from political influence.

Startups and Business Dynamism. One of the key features that characterizes the mechanism behind market power is incomplete passthrough. A firm that operates in a competitive market is unable to influence the price level. Therefore, any shocks to its costs or to the demand translate in an immediate adjustment of the price it charges. When firm's costs fall, it adjusts prices by the amount of the cost decline; if not, it would not sell anything. Likewise, the firm would make losses if it didn't raise prices in the presence of an increase in costs. Prices therefore reflect changes in costs, i.e., cost passthrough to prices is one to one. Instead, when a firm has market power and it can affect the price, the firm's pricing decision is governed by the residual demand elasticity. The firm can raise prices and still sell, and it raises prices to optimally trade off the higher price per unit, with the decline in units sold. The latter is governed by the residual demand elasticity. A change in the cost therefore

[22]The welfare cost of inflation is typically measured as the cost of an increase from 0 to 10% inflation, with different methods obtaining similar estimates, see Lucas (1981), Fischer (1981), Lucas (2000), Cooley and Hansen (1989)

typically results in a less than one to one passthrough adjustment of prices.[23]

Incomplete passthrough has macroeconomic implications for business dynamism and for the startup rate. Decker, Haltiwanger, Jarmin, and Miranda (2020) document that in the aggregate across the economy, the stochastic process of cost or productivity shocks to firms has remained stable. Yet, at the same time, they establish that the firm respond to shocks has become more muted. That means that in the light of the same shocks, firms not respond less. In particular, firms respond less in adjusting the firm size, either output produced or workers employed. The authors measure this responsiveness or dynamism as the reallocation rate, the rate at which firms hire or fire new workers. That reallocation rate has fallen substantially between 1980 and today, despite a similar shock process to which firms are exposed.

This can be rationalized with incomplete passthrough.[24] Firms that have more market power and respond with smaller price changes when shocks occur, will therefore also respond with smaller changes in the quantity produced and thus also in the amount of labor hired in order to produce those quantities. In other words, incomplete passthrough not only manifests itself in prices, it also appears in the volatility of quantities produced and workers employed. The rise of market power in the macroeconomy, via incomplete passthrough, gives rise to a declining reallocation rate.

This decline in business dynamism is not only apparent in the reallocation rate, it also affects the startup rate of new firms. The startup rate is often measured as the fraction of all firms that were founded in the last year (or the last 5 years). That rate has declined from around 14% to 8% (see for example Decker, Haltiwanger, Jarmin, and Miranda (2020), Eeckhout (2021), Akcigit and Ates (2023), Akcigit and Ates (2021), Eggertsson, Mehrotra, and Robbins (2019), Biondi, Inferrera, Mertens, and Miranda (2023)). One explanation points at the role of market power.[25] De Loecker, Eeckhout, and Mongey (2021) show in the context of a model with free entry of firms, that as market power increase – mainly due to a rise in the dispersion of firm productivities and an increase in the fixed cost of entry – there is a decline in the rate of entry and exit, and hence a decline in the startup rate of new firms. The technological change that gives rise to in increase in market power also affects the entry of new firms and changes the size distribution of firms and how it evolves over time. Again, business dynamism is tightly linked to the rise of market power.

Inequality. The rise of economy-wide market power leads to loss of output and a decline in welfare, and at the same time there is a redistribution of the surplus from labor income towards profits. But there is also a rise in inequality within the workforce, between differently skilled workers, a well as

[23] The extent to which and whether this is indeed the case depends on the demand elasticity.

[24] This by no means the only force at work. One alternative explanation has to do with demographics and the composition of the workforce, for example, with an aging workforce and older workers switching jobs less frequently, the changing composition leads to a decline in the aggregate reallocation rate, see for example Dent, Karahan, Pugsley, and Şahin (2016); Engbom (2017); Molloy, Smith, Trezzi, and Wozniak (2016); Hopenhayn, Neira, and Singhania (2022), though Hyatt and Spletzer (2013) argue that demographics alone can only account for a fraction of the decline in the reallocation rate.

[25] There is extensive debate in the economics literature what is driving the decline in the startup rate, and there are several complementary forces at work. One explanation has to do, as in the case of the declining reallocation rate, with demographics and the changing composition of the workforce (see for example Dent, Karahan, Pugsley, and Şahin (2016); Karahan, Pugsley, and Şahin (2016)). Further, Foster, Haltiwanger, and Syverson (2016) show that younger firms have a smaller customer base that takes time to build up. This is consistent with the view that firm size is not exclusively driven by technological factors, but rather by demand factors.

within firms and entrepreneurs.

Deb, Eeckhout, Patel, and Warren (2024) analyze the effect of market power on the skill premium, which measures the ratio between the average wage of workers with college education or more, and the average wage of workers with high school or less. They find a small effect of market power on the skill premium (via market power in the labor market rather than via the final goods market, see also the disussion on monopsony below). Market power predominantly lowers wages, of all education levels alike, and it contributes to an increase in between-establishment wage variance. Market power has also an impact on wage inequality in the top one percent. Bao, De Loecker, and Eeckhout (2023) study the contribution of market power to manager pay. They find that market power contributes on average 45% to the wages of executives. But the impact is modest for managers of small firms, and up to 80% for managers of the top firms. As a result, market power is the driver behind the increase in the inequality in manager pay. Over the period of their sample starting in 1994, manager pay inequality has increased and most of the increase is due to the rise of market power.

Market power has also an impact on the distribution of profits. Earlier in the empirical section, we have discussed the rise of inequality of profits, which is one of the measures of market power. Deb (2023) studies the role of market power for entrepreneurship and shows that income (or profits) of entrepreneurs has increased, yet at the same time the fraction of entrepreneurs has declined. This appears contradictory as one would expect higher earnings to lead to more entry. However, Deb (2023) shows that due to the sharp rise in inequality of entrepreneur earning, the marginal entrepreneur is worse off, even if on average entrepreneurial income has gone up. The increase in inequality of entrepreneurial income is due to to an increase in market power, driven by higher dispersion in firm productivity and an increase in the fixed cost of starting a business. In the context of optimal tax policy, Eeckhout, Fu, Li, and Weng (2025) a similar model and also find a sharp increase in entrepreneurial inequality of income driven by the rise of market power.

Monopsony. In addition to the indirect effect of output market power on wages, there is of course also the direct effect on wages if the labor market is not competitive. Just as in the case of output market power, firms that have market power in the labor market face an upward sloping residual labor supply curve and will as a result set wages and choose their labor force in a way that reflect that pricing power. This is, as Robinson (1932) pointed out, the mirror image of market power in the goods market. In particular, firms pay wages below the marginal product of labor, and higher fewer than the Pareto efficient number of workers.

Robinson coined the phrase 'monopsony' in the labor market to mirror monopoly in the output goods market. Of course, there may be more than one firm, and then the concept is oligopsony (the equivalent of oligopoly). Monopsony and oligopsony of course go beyond the labor market and applies to any input market where firms acquire materials upstream.

The oft-cited example of monopsony is the case of hospital mergers and the effect on the labor market for health care workers. In a medium-sized town with two hospitals, a merger endows the hospital with market power. Prager and Schmitt (2021) find evidence of a drop in wages following the hospital mergers. There is also evidence of a fall in the quality of care (Cooper et al. (2025)). For other recent work see also Rubens (2023),Delabastita and Rubens (2025), Mertens (2019) Mertens and

Mottironi (2023), Morlacco (2017), Azkarate-Askasua and Zerecero (2020).

There has been an interest in monopsony that dates back a long time. The more recent work that aims to put the monopsony problem in an economy-wide setting does so in the context of two classes of models. First, models of labor search exploit the fact that the search friction generates market power to the firm. Manning (2003) builds on the Burdett and Mortensen (1998) model of search with large firms and shows how larger firms have monopsony power over the workers. A second class of models uses the imperfect substitutability between labor inputs, as in Bhaskar and To (1999, 2003). As in the case of models with nested CES technologies of production (see Dixit and Stiglitz (1977); Atkeson and Burstein (2008); Jaimovich and Floetotto (2008)), the monopsony model assumes that labor inputs are imperfect substitutes. In small markets, this leads to market power of firms over the workers.

Recent work uses these settings to quantify the role of labor market power and the effect it has on wages (see for example Berger, Herkenhoff, and Mongey (2022) and Yeh, Macaluso, and Hershbein (2022)). Monopsony or oligopsony power is sizable, with an estimated markdown of 30-50%. However, there is limited evidence of a marked change over the last decades. One open question that has been addressed in Goolsbee and Syverson (2023) and Deb, Eeckhout, Patel, and Warren (2024) is whether monopsony power is stronger or weaker for highly educated workers compared to low education workers. Both papers find that market power is stronger for highly educated workers. Goolsbee and Syverson (2023) analyzes teaching staff at universities and find that those with higher education have higher wages but they face a larger markdown relative to their estimated marginal product. Deb, Eeckhout, Patel, and Warren (2024) finds a similar result using US Census data.

The existence of the direct effect of monopsony power does not mean there cannot be simultaneously indirect, general equilibrium effects from market power in the final goods market on the labor market. Using US Census data, Deb, Eeckhout, Patel, and Warren (2022) asks how much of the widening gap between labor productivity and wages is due to the direct effect of monopsony power and how much is due to the indirect effect of economy-wide output market power that transmits into a declining demand for labor and hence falling wages. They find that 75% is due to the indirect effect of monopoly and 25% is due to the direct effect of monopsony.

Stansbury and Summers (2017) cite yet another reason for the decline in wages relative to productivity. They point to the decline in rent sharing to the detriment of workers, and they point to the falling rate of unionization. This has for an effect that wages get a smaller share of output, and a larger share going to firm profits. This is a valid claim, but in a world with free entry, this begs the question then why, with firms making higher profits, there is not sufficient entry of new competitors which will drive down profits and increase wages. Only if there is some form of market power that prevents firms from entering and obtaining those profits will the change in rent sharing have a lasting impact on wages.

3.2 A Model

We use a simple model of the macroeconomy that builds on earlier work and that incorporates oligopolistic competition in the output markets as in Atkeson and Burstein (2008). The model has

the key determinants of market power, including conduct, which we model in the form of common ownership. We further add returns to scale in production. Here, the presentation of the model, its notation as well as its solution, closely follows Eeckhout (2025). First, we set up the model and then provide the solution.

Environment. A continuum of markets are indexed by $j \in [0, J]$) and there is a finite number of I firms in each market. A firm $i \in [1, ..., I]$ in market j is indexed by i, j.

Preferences. A continuum of representative households supply homogenous labor L. They consume the bundle of all goods C. The household's utility is:

$$U(C, L) = C - \bar{\varphi}^{-\frac{1}{\varphi}} \frac{L_t^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}, \tag{14}$$

where C is the Constant Elasticity of Substitution (CES)[26] aggregator of firm-level consumption C_{ij} and market level consumption C_j, where:

$$C = \left[\int_0^J J^{-\frac{1}{\theta}} c_j^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}, \quad c_j = \left[\sum_{i=1}^{I} I_j^{-\frac{1}{\eta}} c_{ij}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}. \tag{15}$$

θ and η measure the substitutability, with $\theta < \eta$; within-market substitutability is higher than between markets. Households maximize utility by choosing c_{ij} and L constrained by their budget

$$\max_{c_{ij}, L} \; U(C, L) \tag{16}$$

$$\text{s.t} \int_j \sum_i p_{ij} c_{ij} \leq WL + \Pi \tag{17}$$

where the final goods price for i, j is p_{ij} (the sectoral price index is p_j and the aggregate price index is P), and where labor income WL (W is the wage rate) plus a representative portfolio of economy-wide profits Π make up total income.

Technology. Firms produce access a technology with a single input labor l_{ij}, given by $y_{ij} = A_{ij} l_{ij}^{\gamma}$, where γ is the returns-to-scale parameter. The term A_{ij} denotes the firm's Total Factor Productivity (TFP), the log of which follows an AR(1) process, with persistence ρ and innovation ε. Then $a_{ij} = \log A_{ij}$:

$$a_{ij} = \rho a_{ijt-1} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim \mathcal{N}(\mu, \sigma^2). \tag{18}$$

Ensuring the mean of the process A_{ij} is equal to 1, we let $\mu = -\frac{1}{2}\frac{\sigma^2}{1+\rho}$.

[26] Building on the single-nested CES model of Dixit and Stiglitz (1977). Similar models with CES preferences and markups are also widely used in the trade literature, see for example Amiti, Itskhoki, and Konings (2019) and Edmond, Midrigan, and Xu (2015). For a model with a more general preference structure beyond CES, see Pellegrino (2025).

Market Structure. Within each market, firms compete à la Cournot. Their profits are:

$$\pi_{ij} = p_{ij}y_{ij} - W_t l_{ij} - W_t \phi, \tag{19}$$

where $W_t\phi$ is the sunk cost (overhead), expressed in overhead labor units ϕ.

Conduct. To model conduct in a tractable way, we assume that the ownership of firms changes. There is common ownership where firms own also a share $\alpha \in [0,1]$ of all other firms. This highly symmetric formulation is of course a simplification, but it gives a degree of freedom to capture the extent to which conduct changes: see Rotemberg (1984); Azar (2012); Appel, Gormley, and Keim (2016); Ederer and Pellegrino (2025); Azar and Vives (2021)).

The firm's objective is to maximize the $\pi_{ij} + \alpha \sum_{k \neq i} \pi_{kj}$, where α measures the extent to which there is cross-ownership in other firms: $\alpha = 0$ is the standard Cournot solution and $\alpha = 1$ is perfect collusion. Firms solve the following problem:

$$\max_{y_{ij}} \quad p_{ij}y_{ij} - W_t \left(\frac{y_{ij}}{A_{ij}}\right)^{\frac{1}{\gamma}} - W_t\phi_{ij} + \alpha \sum_{k \neq i} \left[p_{kj}y_{kj} - W_t \left(\frac{y_{kj}}{A_{kj}}\right)^{\frac{1}{\gamma}} - W_t\phi_{kj} \right]. \tag{20}$$

We now turn to the solution of the model.

Household solution. Household optimization yields labor supply:

$$L^s(W) = \bar{\varphi}\,(W/P)^{\varphi} \tag{21}$$

as well as the inverse demand function for each good y_{ij} :

$$p_{ij} = I^{-\frac{1}{\eta}} J^{-\frac{1}{\theta}} \left(\frac{y_{ij}}{y_j}\right)^{-\frac{1}{\eta}} \left(\frac{y_j}{Y}\right)^{-\frac{1}{\theta}} P. \tag{22}$$

Firm solution. The firm's optimal solution to (20) satisifes:

$$\frac{\partial \pi_{ij}}{\partial y_{ij}} + \alpha \sum_{k \neq i} \frac{\partial \pi_{kj}}{\partial y_{ij}} = 0 \tag{23}$$

$$p_{ij} + \frac{\partial p_{ij}}{\partial y_{ij}} y_{ij} + \alpha \sum_{k \neq i} \frac{\partial p_{kj}}{\partial y_{ij}} y_{kj} = \frac{W}{A_{ij}^{\frac{1}{\gamma}}} \cdot \frac{1}{\gamma} \cdot y_{ij}^{\frac{1}{\gamma}-1}, \tag{24}$$

where

$$\frac{\partial p_{kj}}{\partial y_{ij}} y_{kj} = \left(\frac{1}{\eta} - \frac{1}{\theta}\right) s_{ij} \frac{s_{kj}}{s_{ij}} p_{ij} \tag{25}$$

with s_{ij} is the firm's sales share. The FOC can be written as:

$$p_{ij}\left[1-\frac{1}{\eta}+\left(\frac{1}{\eta}-\frac{1}{\theta}\right)\left(s_{ij}+\alpha\sum_{k\neq i}s_{kj}\right)\right]=\frac{W}{A_{ij}^{\frac{1}{\gamma}}}\cdot\frac{1}{\gamma}\cdot y_{ij}^{\frac{1}{\gamma}-1}\quad\Leftrightarrow\quad p_{ij}\underbrace{\left(1+\varepsilon_{ij}^{p}\right)}_{\mu_{ij}^{-1}}=MC_{ij}\tag{26}$$

with MC_{ij} is the marginal cost, and ε_{ij}^{p} the residual demand elasticity. The markup is equal to:

$$\mu_{ij}=\frac{p_{ij}}{MC_{ij}}=\left[1-\frac{1}{\eta}+\left(\frac{1}{\eta}-\frac{1}{\theta}\right)\left(\alpha+s_{ij}(1-\alpha)\right)\right]^{-1}.\tag{27}$$

The sales shares s_{ij} implicitly solves the fixed point to

$$s_{ij}=\frac{\left[\frac{1}{\left(1-\frac{1}{\eta}+\left(\frac{1}{\eta}-\frac{1}{\theta}\right)\left(\alpha+s_{ij}(1-\alpha)\right)\right)}\times A_{ij}^{-\frac{1}{\gamma}}\right]^{\frac{\gamma-\gamma\eta}{\gamma-\gamma\eta+\eta}}}{\sum_{k=1}^{I}\left[\frac{1}{\left(1-\frac{1}{\eta}+\left(\frac{1}{\eta}-\frac{1}{\theta}\right)\left(\alpha+s_{kj}(1-\alpha)\right)\right)}\times A_{kj}^{-\frac{1}{\gamma}}\right]^{\frac{\gamma-\gamma\eta}{\gamma-\gamma\eta+\eta}}}=\frac{\left(\frac{\mu_{ij}}{A_{ij}^{1/\gamma}}\right)^{\frac{\gamma-\gamma\eta}{\gamma-\gamma\eta+\eta}}}{\sum_{k}\left(\frac{\mu_{kj}}{A_{kj}^{1/\gamma}}\right)^{\frac{\gamma-\gamma\eta}{\gamma-\gamma\eta+\eta}}}.\tag{28}$$

We derive the prices in the Online Appendix A.1.

Labor demand. A firm's individual variable labor demand is $l_{ij}=\left(\frac{y_{ij}}{A_{ij}}\right)^{1/\gamma}$, which aggregates to, using the equilibrium quantities y_{ij}:

$$L^{v}=\int_{j}\sum_{i}l_{ij}\,dj=\int_{j}\sum_{i}\left(\frac{y_{ij}}{A_{ij}}\right)^{\frac{1}{\gamma}}dj=Y^{\frac{1}{\gamma}}J^{-\frac{1}{\gamma}}\int_{j}I^{\frac{\theta-1}{\gamma(1-\eta)}}\sum_{i}\left(\frac{s_{ij}^{\frac{\theta-\eta}{1-\eta}}p_{ij}^{-\theta}P^{\theta}}{A_{ij}}\right)^{\frac{1}{\gamma}}dj\tag{29}$$

Firms also employ a fixed amount of labor for overhead, which in the aggregate is $L^{f}=\int_{j\in J}\sum_{i\in I}\phi_{ij}$. Then total labor demand is $L^{d}=L^{f}+L^{v}$. Labor-market clearing, equates demand and supply $L^{d}=L^{s}(W)=\bar{\varphi}(W/P)^{\varphi}$, and pins down the wage W once the aggregates Y and P have been determined from prices.

Market equilibrium. We normalize the aggregate price index $P=1$. The economy is then characterised by three economy-wide conditions that jointly determine the wage W, aggregate output Y, and total employment L, satisfying labor demand, labor supply and goods market clearing. In the Online Appendix A.2 we lay out the solution algorithm of the economy.

3.3 Quantification and Comparative Statics

We choose parameters based on earlier work that approximate the properties to what we observe for the US economy in the last four decades. We report the baseline parameters in Table 1. The objective

is to illustrate how the main properties of the economy change when the primitive determinants change.

Table 1: Baseline Parameters

Within-market demand elasticity	η	5.75	De Loecker, Eeckhout, and Mongey (2021)
Between-market demand elasticity	θ	1.25	De Loecker, Eeckhout, and Mongey (2021)
Labor supply elasticity	φ	0.25	Chetty, Guren, Manoli, and Weber (2011)
Labor supply intercept	$\bar{\varphi}$	8×10^6	Compustat
Technology Dispersion	σ	1	Exogenously set
Returns-to-scale parameter	γ	1	Exogenously set
Overhead cost per firm	ϕ	100	Exogenously set
Common-ownership	α	0.25	Exogenously set
Number of markets	J	1000	Exogenously set
Number of firms in each market	I	16	Exogenously set

We investigate the impact of technology (dispersion of shocks and returns to scale), demand (elasticity of substitution), market structure (number of competitors) and conduct (extent of common ownership). We measure the impact of these parameters on aggregate markups; the distribution of markups; the profit rate; labor income, output and welfare; the labor share; and the variance of markups measured by the percentage difference between the revenue and cost-weighted markups.

In the Appendix[27] we perform comparative statics on each of the model parameters and evaluate how each parameter affects the economy in terms of markups, profits, output, labor income and welfare.

3.4 Aggregate Outcomes when Aggregate Markups are Constant

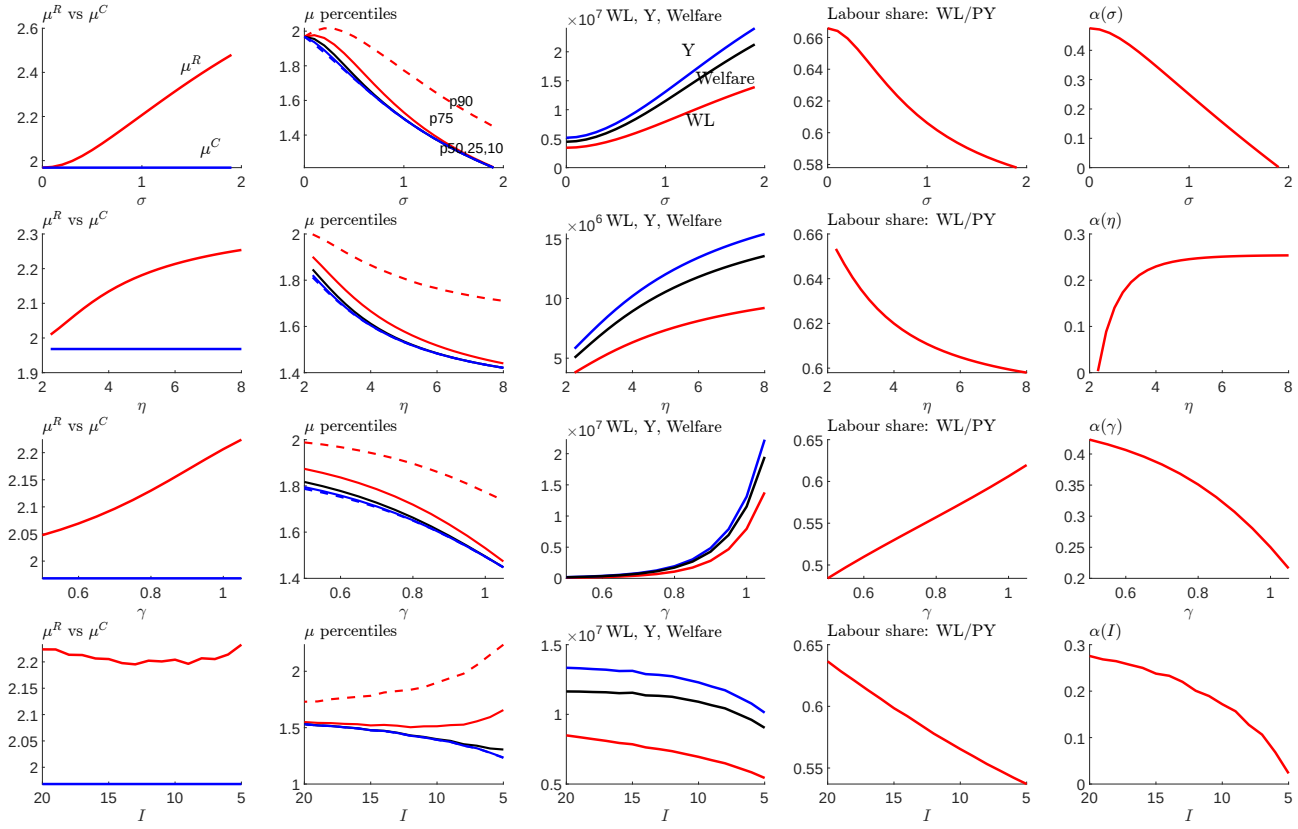
To illustrate the point that aggregate markups alone do not pin down welfare, we perform a simple exercise on our baseline model. We fix the cost-weighted aggregate markup μ^C, and we measure the effect of each of the parameters on the key endogenous outcomes, most notably welfare, but also output, labor income, markups,... Key is that in order to keep μ^C constant, we need to adjust at least one parameter.

In Figure 3a, we perform this comparative statics exercise. We adjust the conduct parameter α to keep μ^C constant. We then vary 4 of the other parameters each in turn: σ, η, γ, and I. In the last column, we plot how α must adjust when each of the parameters changes in order to keep μ_C constant: α is decreasing in σ and γ, and increasing in η and I. The effect of σ and γ is to increase the revenue-weighted markup. This comes from two effects. First, the direct effect from σ, γ is to increase the dispersion of revenue which affects the revenue-weighted markup much more than the increase in the cost-weights. Second, the indirect effect from adjusting α in order to keep μ^C constant: α declines which lowers the markups and offsets the increased dispersion in costs. The dispersion in revenue increases more, leading to the rise in revenue-weighted markups. The increased dispersion

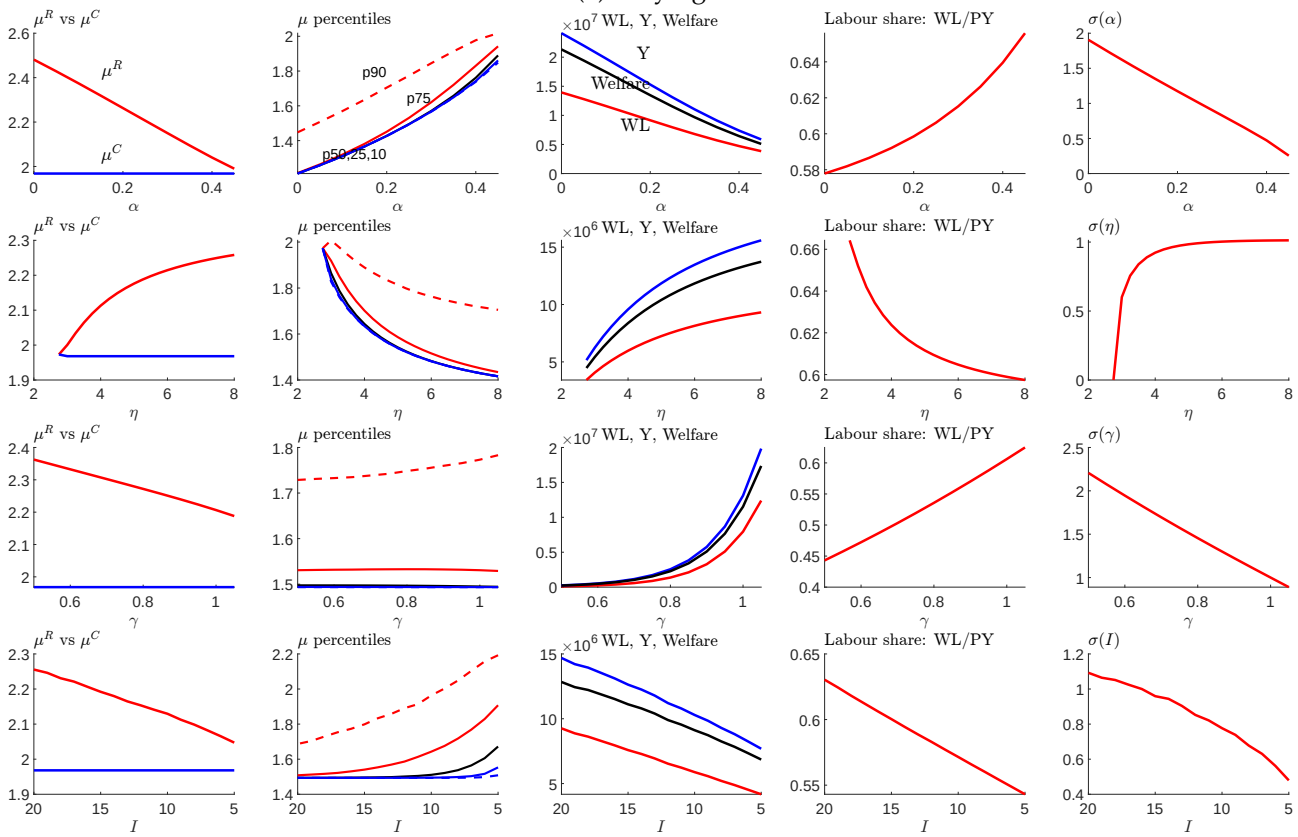
[27] See Figures B.12,B.13,B.14,B.15,B.16,B.17 in the Appendix.

from σ and γ increases output, the wage bill and welfare. But it has substantial distributional effects as the labor share drops.

The direct effect of η is to reduce markups, so to keep μ^C constant, we need an increase in α. This in turn affects the revenue distribution more than the cost distribution as market power increases. It leads nonetheless to an increase in welfare driven by utility, but again, with redistribution and a drop in the labor share. Finally, market structure I has virtually no effect on μ^R. As the number of firms drops, markups increase but they also become more dispersed. To keep μ^C constant, we need a drop in the α. The net effect is a drop in welfare and output, and redistribution with a drop in the labor share and thus an increase in the profit share.



(a) Varying α



(b) Varying σ

Figure 3: Comparative statics to maintain a constant cost-weighted markup.

Notes. We first compute the cost-weighted markup for the benchmark case defined by the parameter set in Table 1, then vary one parameter at a time and solve for the value of α that keeps the cost-weighted markup constant. Finally, we plot the resulting comparative statics.

We repeat similar exercises to keep μ^C constant but where sequentially we change σ, η and γ, reported in Figures B.18, B.19, and B.20 in the Appendix. The main implications are the same. As we keep μ^C constant by varying $\sigma, \eta.\gamma$, while evaluating the effect of the other parameters, we find that firms adjust their behavior, which results in changing distributions of costs, revenue and markups, and hence the resulting differences in aggregate markups and welfare and output.

The main insight from these exercises is that an aggregate markup alone cannot inform us about welfare, output, the labor share or any other macroeconomic outcomes. All determinants – preferences, conduct, market structure, technology – play a role, and any change in the underlying parameters results in a change in the *behavior* of the firms and how they compete. This automatically results in changes in the underlying outcome variables: markups, output, wages and welfare. The key determinant here is the distribution, especially when focusing on aggregate markups. We cannot reach any conclusions about the extent of market power from observing aggregate markups without knowledge of the underlying joint distribution of parameters (technology in particular), and the resulting distributions of revenue, costs and prices.

4 Competition Policy as Redistribution

When there is market power, the first theorem of welfare economics fails. Productive resources are not efficiently allocated, and part of the surplus is lost. When market power is economy-wide, within the context of a model we can duly define welfare taking into account the general equilibrium effects via prices and wages. The first observation is that welfare is much broader than merely the consumer surplus, which is a partial equilibrium notion. Instead we calculate welfare from aggregating utilities across all agents in the economy, taking into account that households are the owners of the firms. Of course, the social welfare criterion chosen to aggregate those individual utilities crucially determines policy, but this is something economists in public finance are routinely dealing with. An important caveat here is that there are of course sources of inefficiencies and frictions other than market power. Here we discuss policy solutions as if market power is the only friction in the economy. Policy in the real world deals with multiple frictions simultaneously.

The key feature of the rise of market power that we have discussed throughout this review article is heterogeneity. The causes of market power stem from differences across different firms. This creates inequalities between firms: not all firms gain from the rise of market power. And there are plenty of other forms of inequality that result from market power. In fact, market power is really all about distribution: between wages for workers and profits for firms; between profits for dominant firms and profits for small startups; between wages of superstars and wages of production and service workers. If market power exacerbates those inequalities, the remedying market power will reduce them. Competition policy therefore not only restores efficiency, it is also about the second welfare theorem, namely redistribution by means of the market mechanism. Competition policy as redistribution acknowledges that by restoring inefficiencies due to market power, we also reduce the inequalities associated with it.

In most countries, reducing inequality is not explicitly in the mandate of competition authorities

(in merger review for example). What we want to point out here is that competition policy does have an impact on the distribution of resources, whether that is intended with a mandate or not. The first and foremost policy to remedy the negative effects of market power is therefore competition policy, which we discuss first. Then we continue with discussing a non-exhaustive list of other policies beyond competition policy, and what the implications are for welfare.

Competition policy. Competition law in the US and in most advanced economies is broadly defined. In practice, competition policy is often a lot narrower with a focus on anti-competitive conduct such as collusion and cartels, abuse of dominant positions and mergers that harm competition. For example, competition authorities rightly address harm from price fixing, predatory pricing and dominance from mergers. Despite the vigilance of competition authorities, we have nonetheless seen a substantial increase in all measures of market power. All competition agencies work with severely limited resources and can review a limited number of cases.

More importantly, there continues to be a focus on explicit actions that show abuse of dominance. The facts that we have discussed earlier show that the technological change in the digital age is characterized by growing increasing returns to scale and a rise in economies scale. It is hard to establish that a firm that operates a technology with increasing returns is abusing its dominant position.

Market power where dominant positions stem from technologies that lead to natural monopoly is often addressed with ex ante regulation rather than ex post enforcement. Historically, most advanced economies address the scale issue that arises in sectors such as utilities (gas, water, electricity,...), transportation, natural resources and telecomunications with regulation to ensure competitive prices. This is achieved with varying degrees of success and forms of implementation.

Interoperability is a way to achieve such competition. For example, the regulator can achieve competitive pricing on high speed railway lines, a natural monopoly, by making the technology interoperable. The regulator separates the rail network from the operators of the train. Operators have access to the network at user price that is set by the regulator. If the access price is set low enough there will be multiple entrants who compete for the customers. If the price is too high there are not enough resources to invest in maintenance and innovation of the network. If one firm owns both the network and is the sole operator, it has no incentive to allow competitors to access the network, or equivalently, it will want to set a prohibitively high access price. Interoperability is most widely used in telecommunications, where different operators have access to the same network of phone numbers for example, or where users of competing Internet Service Providers can send email messages to each other across competitors. Such interoperability is not imposed in markets that use more recent technology, such as messaging for example. You cannot send a message from a WhatsApp account to a Telegram account, for example. In the European Union, the Digital Markets Act (DMA) is recent legislation that aims to introduce competition using tools such as interoperability. In the absence of effective competition policy, we now discuss the impact of some conventional policies that aim to remediate the negative impact of market power.

Taxation. Income and profit taxation is a tool widely suggested to alleviate the negative consequences of market power. Taxation can be both efficiency-correcting and redistributive. Unfortu-

nately, taxation cannot restore first-best. To see this, consider the simplest partial equilibrium setting where profits are independent of the effort of the entrepreneur and wages are fixed. Then taxing profits and subsidizing labor income will redistribute from entrepreneurs to households, but it will not improve efficiency. Firms still charge high markups and as a result of the high prices, fewer customers buy the products than under competition. This leads to a welfare loss, even if the gains of those high markups are redistributed to the workers.

Market power has an externality, and the optimal tax is a Pigouvian correction for this externality. The planner wants the tax to overcome that externality. Paradoxically, a tax subsidizes production because market power leads to under-production, and a subsidy that entices firms to produce more can restore Pareto efficiency. Without lump sum transfers that further increases inequality of course.

In the presence of incomplete information, the Mirrleesian income tax policy takes into account not only the Pigouvian correction of the externality of market power, but also the incentive effect to provide effort. Now the optimal tax provides incentives for heterogenously skilled workers to provide effort. In a world with an incentive problem for the workers, Boar and Midrigan (2024) find that the subsidy proportional to the market power and that corrects the Pigouvian externality dominates. However, once we take into account an incentive problem for the entrepreneurs – arguably the most relevant incentive problem – Eeckhout, Fu, Li, and Weng (2025) show that optimal policy corrects the rise of market power with higher profit taxes and lower labor income taxes. This optimal tax policy reduces the inefficiency of market power and redistributes resources from profits to labor income.

Minimum Wages. Researchers and policy makers have proposed minimum wages as a way to deal with a falling labor share and a widening wage-productivity gap. The empirical work by Card and Krueger (1994) refutes the view based on competitive markets that higher minimum wages reduce employment. Stigler (1946) shows that in competitive markets, a minimum wage higher than the equilibrium wage leads to a reduction in employment and efficiency. However, using data on a case study in the labor market in the fast food industry, Card and Krueger (1994) show that in that industry higher minimum wages do not affect employment.

Recent research, modeling the labor market in a macroeconomic setting with market power confirms the Card and Krueger (1994) finding. As long as the minimum wage is not too much above the equilibrium wage, the minimum wage does not lead to a decrease in employment and increases welfare (see Berger, Herkenhoff, and Mongey (2025) and Lo Bello and Pesaresi (2022) for example). In fact, for minimum wage levels close to the equilibrium wage may lead to an increase in employment. Eventually, when the minimum wage level is too high, there is a decline in employment. This effect is driven by the presence of monopsony power. The minimum wage changes the (inefficient) first-order condition and for low levels of the minimum wage does not decline employment. This effect is not present when there is only output market power and no monopsony power.

Maximum wage caps for executives. Some policy proposals recommend a maximum wage cap on executive compensation. As manager wages have grown substantially, a maximum wage cap is seen as a way to reign in wage excesses. In government agencies in particular there are often maximum wage restrictions. In a setting where wages of managers reflect in part the contribution

from market power, Bao et al. (2023) analyze the impact of a maximum wage cap. It turns out that in that setting, managers sort into heterogenous firms. The maximum wage gap distorts the allocation mechanism, which under unrestricted wages ensures that the most talented manager ends up in the most productive firm. With a maximum wage cap that is binding, i.e. below the highest wage, this allocation mechanism is no longer at work for all mangers for whom the maximum wage is binding. Bao et al. (2023) show that while this compresses wages, this also leads to large output losses due to the misallocation of managers.

5 Conclusion

In this paper we have reviewed some of the recent developments of measurement and interpretation of market power in the macroeconomy. We confirm the evolution of economy-wide market power using different measures. Markups, the most widely used measure, are robust across different methods such as production function estimation or cost shares. We find that the rise of markups since 1980 is driven both by reallocation towards high markup firms and by the increase of the revenue to cost ratio. Across different methods, the rise of markups is not due to the change in output elasticities.

In the context of a model economy, we evaluate the welfare implications of market power. We find that welfare effects crucially depend the distributional characteristics in the economy, and welfare cannot be summarized by one aggregate measure alone.

References

ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2012): "The network origins of aggregate fluctuations," *Econometrica*, 80, 1977–2016.

ACKERBERG, D., C. L. BENKARD, S. BERRY, AND A. PAKES (2007): "Econometric tools for analyzing market outcomes," *Handbook of econometrics*, 6, 4171–4276.

ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): "Identification properties of recent production function estimators," *Econometrica*, 83, 2411–2451.

AGHION, P., A. BERGEAUD, T. BOPPART, P. J. KLENOW, AND H. LI (2023): "A theory of falling growth and rising rents," *Review of Economic Studies*, 90, 2675–2702.

AKCIGIT, U. AND S. T. ATES (2021): "Ten facts on declining business dynamism and lessons from endogenous growth theory," *American Economic Journal: Macroeconomics*, 13, 257–298.

——— (2023): "What happened to US business dynamism?" *Journal of Political Economy*, 131, 2059–2124.

AMITI, M., O. ITSKHOKI, AND J. KONINGS (2019): "International shocks, variable markups, and domestic prices," *The Review of Economic Studies*, 86, 2356–2402.

APPEL, I. R., T. A. GORMLEY, AND D. B. KEIM (2016): "Passive investors, not passive owners," *Journal of Financial Economics*, 121, 111–141.

ATKESON, A. AND A. BURSTEIN (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98, 1998–2031.

AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2020): "The Fall of the Labor Share and the Rise of Superstar Firms," *Quarterly Journal of Economics*, 135, 645–709.

AZAR, J. (2012): "A new look at oligopoly: Implicit collusion through portfolio diversification," Ph.D. thesis, Princeton University.

AZAR, J. AND X. VIVES (2021): "General equilibrium oligopoly and ownership structure," *Econometrica*, 89, 999–1048.

AZKARATE-ASKASUA, M. AND M. ZERECERO (2020): "The aggregate effects of labor market concentration," *Unpublished Working Paper*.

BAO, R., J. DE LOECKER, AND J. EECKHOUT (2023): "Manager Pay Inequality and Market Power," UPF mimeo.

BARKAI, S. (2020): "Declining Labor and Capital Shares," *Journal of Finance, Forthcoming*, 75, 2421–2463.

BASU, S. (2019): "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence," *Journal of Economic Perspectives*, 33, 3–22.

BERGER, D., K. HERKENHOFF, AND S. MONGEY (2022): "Labor market power," *American Economic Review*, 112, 1147–1193.

——— (2025): "Minimum Wages, Efficiency, and Welfare," *Econometrica*, 93, 265–301.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile prices in market equilibrium," *Econometrica*, 841–890.

BHASKAR, V. AND T. TO (1999): "Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition," *The Economic Journal*, 109, 190–203.

——— (2003): "Oligopsony and the distribution of wages," *European Economic Review*, 47, 371–399.

BIONDI, F., S. INFERRERA, M. MERTENS, AND J. MIRANDA (2023): "Declining business dynamism in Europe," Tech. rep., Jena Economic Research Papers.

BOAR, C. AND V. MIDRIGAN (2024): "Markups and inequality," *Review of Economic Studies*, rdae103.

BOND, S., A. HASHEMI, G. KAPLAN, AND P. ZOCH (2021): "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data," *Journal of Monetary Economics*, 121, 1–14.

BRAND, J. (2021): "Differences in differentiation: Rising variety and markups in retail food stores," Microsoft mimeo.

BRESNAHAN, T. F. (1989): "Empirical studies of industries with market power," *Handbook of industrial organization*, 2, 1011–1057.

BURDETT, K. AND D. T. MORTENSEN (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 257–273.

CARD, D. AND A. B. KRUEGER (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *The American Economic Review*, 84, 772–793.

CASAL, L. (2025): "Lock-In and Productive Innovations," Cornell mimeo.

CHETTY, R., A. GUREN, D. MANOLI, AND A. WEBER (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–75.

CHIAVARI, A. (2024): "The Macroeconomics of Rising Returns to Scale," Oxford mimeo.

COLLARD-WEXLER, A. AND J. DE LOECKER (2015): "Reallocation and Technology: Evidence from the US Steel Industry," *The American Economic Review*, 105.

COOLEY, T. F. AND G. D. HANSEN (1989): "The inflation tax in a real business cycle model," *The American Economic Review*, 733–748.

COOPER, Z., S. V. CRAIG, A. EPANOMERITAKIS, M. GRENNAN, J. R. MARTINEZ, F. S. MORTON, AND A. T. SWANSON (2025): "Are Hospital Acquisitions of Physician Practices Anticompetitive?" Tech. rep., National Bureau of Economic Research.

DE LOECKER, J. (2011): "Recovering markups from production data," *International Journal of Industrial Organization*, 29, 350–355.

DE LOECKER, J. AND J. EECKHOUT (2017): "The Rise of Market Power and the Macroeconomic Implications," NBER working paper 23687.

DE LOECKER, J., J. EECKHOUT, AND S. MONGEY (2021): "Quantifying market power and business dynamism in the macroeconomy," Tech. rep., National Bureau of Economic Research.

DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 135, 561–644.

DE LOECKER, J. AND P. SCOTT (2025): "Markup estimation using production and demand data. An application to the us brewing industry," *Review of Economic Studies*, Forthcoming.

DE LOECKER, J. AND C. SYVERSON (2021): "An Industrial Organization Perspective on Productivity," *Handbook of Industrial Organization*, IV.

DE LOECKER, J. AND F. M. P. WARZYNSKI (2012): "Markups and Firm-level Export Status," *American Economic Review*, 102, 2437–2471.

DE RIDDER, M. (2024): "Market power and innovation in the intangible economy," *American Economic Review*, 114, 199–251.

DE RIDDER, M., B. GRASSI, AND G. MORZENTI (2025, forthcoming): "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates from Financial Data," *Econometrica*.

DEB, S. (2023): "How Market Structure Shapes Entrepreneurship and Inequality," UPF mimeo.

DEB, S., J. EECKHOUT, A. PATEL, AND L. WARREN (2022): "What drives wage stagnation: Monopsony or Monopoly?" *Journal of the European Economic Association*, 20, 2181–2225.

——— (2024): "Market Power and Wage Inequality," *Econometrica*, forthcoming.

DECKER, R. A., J. HALTIWANGER, R. S. JARMIN, AND J. MIRANDA (2020): "Changing business dynamism and productivity: Shocks versus responsiveness," *American Economic Review*, 110, 3952–3990.

DELABASTITA, V. AND M. RUBENS (2025): "Colluding against workers," *Journal of Political Economy*, 133, 1796–1839.

DENT, R. C., F. KARAHAN, B. PUGSLEY, AND A. ŞAHIN (2016): "The Role of Startups in Structural Transformation," *American Economic Review*, 106, 219–23.

DIXIT, A. K. AND J. E. STIGLITZ (1977): "Monopolistic competition and optimum product diversity," *The American economic review*, 67, 297–308.

DÖPPER, H., A. MACKAY, N. H. MILLER, AND J. STIEBALE (2025): "Rising markups and the role of consumer preferences," *Journal of Political Economy*, 133, 000–000.

EDERER, F. AND B. PELLEGRINO (2025): "A tale of two networks: Common ownership and product market rivalry," *Review of Economic Studies*, forthcoming.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2015): "Competition, markups, and the gains from international trade," *American Economic Review*, 105, 3183–3221.

——— (2023): "How costly are markups?" *Journal of Political Economy*, 131, 1619–1675.

EECKHOUT, J. (2020): "Comment on: Diverging Trends in National and Local Concentration," in *NBER Macroeconomics Annual 2020, Volume 35*, NBER.

——— (2021): *The Profit Paradox. How Thriving Firms Threaten the Future of Work*, Princeton, NJ: Princeton University Press.

——— (2025): "The Value and Profits of Firms," *Journal of the European Economic Association*, 23, 52–118.

EECKHOUT, J., C. FU, W. LI, AND X. WENG (2025): "Optimal Taxation and Market Power," *American Economic Review*, forthcoming.

EECKHOUT, J. AND L. VELDKAMP (2022): "Data and Market Power," Columbia mimeo.

EGGERTSSON, G. B., N. R. MEHROTRA, AND J. A. ROBBINS (2019): "A model of secular stagnation: Theory and quantitative evaluation," *American Economic Journal: Macroeconomics*, 11, 1–48.

ENGBOM, N. (2017): "Firm and Worker Dynamics in an Aging Labor Market," Tech. rep., Working Paper.

FISCHER, S. (1981): "Towards an Understanding of the Costs of Inflation: II," in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 15, 5–41.

FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *American Economic Review*, 98, 394–425.

——— (2016): "The slow growth of new plants: Learning about demand?" *Economica*, 83, 91–129.

FOSTER, L., J. HALTIWANGER, AND C. TUTTLE (2024): "Rising Markups or Changing Technology?" Mimeo University of Maryland.

GABAIX, X. (2011): "The granular origins of aggregate fluctuations," *Econometrica*, 79, 733–772.

GANDHI, A., S. NAVARRO, AND D. A. RIVERS (2020): "On the identification of gross output production functions," *Journal of Political Economy*, 128, 2973–3016.

GOOLSBEE, A. AND C. SYVERSON (2023): "Monopsony Power in Higher Education: A Tale of Two Tracks," *Journal of Labor Economics*, 41.

GRASSI, B. (2017): "IO in IO: Competition and Volatility in Input-Output Networks," *Unpublished Manuscript, Bocconi University*.

GRIECO, P., C. MURRY, AND A. YURUKOGLU (2024): "Evolution of Market Power in the US Auto Industry," *Quarterly Journal of Economics*, 139.

GRULLON, G., Y. LARKIN, AND R. MICHAELY (2019): "Are US industries becoming more concentrated?" *Review of Finance*, 23, 697–743.

HALL, R. (1988): "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96, 921–947.

HALL, R. E. (2018): "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy," Tech. rep., National Bureau of Economic Research.

HALTIWANGER, J. C. (1997): "Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence," *Federal Reserve Bank St. Louis Review*, 79, 55–77.

HANNAH, L. AND J. A. KAY (1977): *Concentration in modern industry: Theory, measurement and the UK experience*, Springer.

HOPENHAYN, H., J. NEIRA, AND R. SINGHANIA (2022): "From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share," *Econometrica*, 90, 1879–1914.

HUBMER, J., M. CHAN, S. OZKAN, S. SALGADO, AND G. HONG (2024): *Scalable versus Productive Technologies*, Penn Institute for Economic Research, Department of Economics, University of ....

HYATT, H. R. AND J. R. SPLETZER (2013): "The recent decline in employment dynamics," *IZA Journal of Labor Economics*, 2, 5.

JAIMOVICH, N. AND M. FLOETOTTO (2008): "Firm dynamics, markup variations, and the business cycle," *Journal of monetary Economics*, 55, 1238–1252.

KARABARBOUNIS, L. AND B. NEIMAN (2014): "The Global Decline of the Labor Share," *Quarterly Journal of Economics*, 129.

KARAHAN, F., B. PUGSLEY, AND A. ŞAHIN (2016): "Demographic Origins of the Startup Deficit," NY Fed mimeo.

KEHRIG, M. AND N. VINCENT (2021): "The micro-level anatomy of the labor share decline," *The Quarterly Journal of Economics*, 136, 1031–1087.

KLETTE, T. J. AND Z. GRILICHES (1996): "The inconsistency of common scale estimators when output prices are unobserved and endogenous," *Journal of applied econometrics*, 11, 343–361.

LO BELLO, S. AND L. PESARESI (2022): "Equilibrium Effects of the Minimum Wage: The Role of Product Market Power," Bank of Italy mimeo.

LUCAS, R. E. (1981): "Discussion of: Stanley Fischer,"Towards an understanding of the costs of inflation: II"," in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 15, 43–52.

LUCAS, JR, R. E. (2000): "Inflation and welfare," *Econometrica*, 68, 247–274.

MANNING, A. (2003): *Monopsony in motion: Imperfect competition in labor markets*, Princeton University Press.

MERTENS, M. (2019): "Micro-Mechanisms behind Declining Labor Shares: Market Power, Production Processes, and Global Competition," Halle Institute for Economic Research mimeo.

MERTENS, M. AND B. MOTTIRONI (2023): "Do larger firms have higher markups?" Tech. rep., IWH Discussion Papers.

MILLER, N. H., M. OSBORNE, G. SHEU, AND G. SILEO (2025): "Technology and Market Power in the Cement Industry," Georgetown University.

MOLLOY, R., C. SMITH, R. TREZZI, AND A. WOZNIAK (2016): "Understanding Declining Fluidity in the U.S. Labor Market," *Brookings Papers on Economic Activity*.

MORLACCO, M. (2017): "Market Power in Input Markets: Theory and Evidence from French Manufacturing," Tech. rep., Yale.

OLLEY, G. S. AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1263–1297.

OLMSTEAD-RUMSEY, J. (2019): "Market concentration and the productivity slowdown," *MPRA paper*, 93260.

PELLEGRINO, B. (2025): "Product differentiation and oligopoly: A network approach," *American Economic Review*, 115, 1170–1225.

PRAGER, E. AND M. SCHMITT (2021): "Employer consolidation and wages: Evidence from hospitals," *American Economic Review*, 111, 397–427.

ROBINSON, J. (1932): *The economics of imperfect competition*, Springer.

ROSSI-HANSBERG, E., P.-D. SARTE, AND N. TRACHTER (2021): "Diverging trends in national and local concentration," *NBER Macroeconomics Annual*, 35, 115–150.

ROTEMBERG, J. (1984): "Financial transaction costs and industrial performance," Tech. rep., MIT Sloan working paper.

RUBENS, M. (2023): "Market structure, oligopsony power, and productivity," *American Economic Review*, 113, 2382–2410.

STANSBURY, A. M. AND L. H. SUMMERS (2017): "Productivity and Pay: Is the link broken?" Working paper, National Bureau of Economic Research.

STIGLER, G. J. (1946): "The economics of minimum wage legislation," *The American Economic Review*, 36, 358–365.

SUTTON, J. (1991): *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration*, MIT press.

——— (2001): *Technology and market structure: theory and history*, MIT press.

SYVERSON, C. (2004): "Market structure and productivity: A concrete example," *Journal of political Economy*, 112, 1181–1222.

——— (2019): "Macroeconomics and Market Power: Context, Implications, and Open Questions," *Journal of Economic Perspectives*, 33, 23–43.

TRAINA, J. (2018): "Is Aggregate Market Power Increasing? Production Trends using Financial Statements," Chicago Booth mimeo.

YEH, C., C. MACALUSO, AND B. HERSHBEIN (2022): "Monopsony in the US labor market," *American Economic Review*, 112, 2099–2138.

Online Appendix

A Model

A.1 Omitted Derivations

Prices We can rewrite prices as a function of aggregate ouput:

$$p_{ij} = \left\{ \frac{W}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} P^{\theta} Y\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma}{\gamma-\gamma\theta+\theta}} \tag{A.1}$$

and aggregate at industry and market level using the nested CES structure:

$$p_j = \left\{ \frac{1}{I} \sum_{i=1}^{I} \left\{ \frac{W}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} P^{\theta} Y\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma(1-\eta)}{\gamma-\gamma\theta+\theta}} \right\}^{\frac{1}{1-\eta}} \tag{A.2}$$

$$P = \left\{ \frac{1}{J} \int \left\{ \frac{1}{I} \sum_{i=1}^{I} \left\{ \frac{W}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} P^{\theta} Y\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma(1-\eta)}{\gamma-\gamma\theta+\theta}} \right\}^{\frac{1-\theta}{1-\eta}} \right\}^{\frac{1}{1-\theta}}. \tag{A.3}$$

We normalize the aggregate price level to 1, and determine aggregate wages W and total output Y from market clearing.

We can easily determine the firm-level prices solving first for the combination of wages and aggregate output $W^{\gamma} Y^{1-\gamma}$. Once $W^{\gamma} Y^{1-\gamma}$ is known, the firm-level prices follow mechanically from the formula above:

$$p_{ij} = \frac{\left\{ \frac{1}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} P^{\theta}\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma}{\gamma-\gamma\theta+\theta}}}{\left\{ \frac{1}{J} \int \left\{ \frac{1}{I} \sum_{i=1}^{I} \left\{ \frac{1}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}}\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma(1-\eta)}{\gamma-\gamma\theta+\theta}} \right\}^{\frac{\theta-1}{\eta-1}} \right\}^{\frac{1}{1-\theta}}}$$

and sector and aggregate prices are obtained by the CES aggregations.

Special case when $\gamma = 1$ Normalize the aggregate price index $P = 1$ to be the numeraire, and one can immediately solve for the wage after obtaining sales shares and markups. Goods market clearing yields the equilibrium wage, which can be expressed as:

$$\frac{W}{P} = \left[\int_j \frac{1}{J} \left(\sum_i \frac{1}{I} \left(\frac{\mu_{ij}}{A_{ij}} \right)^{1-\eta} \right)^{\frac{1-\theta}{1-\eta}} \right]^{-\frac{1}{1-\theta}}. \tag{A.4}$$

Labor supply satisfies:

$$L_s = \bar{\varphi} \left(\frac{W}{P} \right)^{\varphi}, \tag{A.5}$$

and the labor market clearing condition:

$$L_{prod} + L_{overhead} = L_s \tag{A.6}$$

where labor demand, $L_{prod} = \int_j \sum_i l_{ij}$, $L_{overhead} = \int_j \sum_i \phi$.

A.2 Solution Algorithm

(i) **Solve the block-recursive micro block.**

1. For each market j, solve the fixed-point system

$$s_{ij} = \frac{\left[\frac{1}{\left(1 - \frac{1}{\eta} + \left(\frac{1}{\eta} - \frac{1}{\theta}\right)\left(\alpha + s_{ij}(1-\alpha)\right)\right)} \times A_{ij}^{-\frac{1}{\gamma}} \right]^{\frac{\gamma-\gamma\eta}{\gamma-\gamma\eta+\eta}}}{\sum_{k=1}^{I} \left[\frac{1}{\left(1 - \frac{1}{\eta} + \left(\frac{1}{\eta} - \frac{1}{\theta}\right)\left(\alpha + s_{kj}(1-\alpha)\right)\right)} \times A_{kj}^{-\frac{1}{\gamma}} \right]^{\frac{\gamma-\gamma\eta}{\gamma-\gamma\eta+\eta}}}$$

   to obtain the sales shares $\{s_{ij}\}$.

2. Compute firm-level markups as

$$\mu_{ij} = \left[1 - \frac{1}{\eta} + \left(\frac{1}{\eta} - \frac{1}{\theta} \right) \left(\alpha + s_{ij}(1-\alpha) \right) \right]^{-1}.$$

(ii) **Back-out the firm-level prices.**

1. Using s_{ij} and μ_{ij}, compute firm-level prices $W^{\gamma} Y^{1-\gamma}$ from the price-index condition:

$$\underbrace{W^{\gamma} Y^{1-\gamma}}_{W_Y} = \left\{ \frac{1}{J} \int \left\{ \frac{1}{I} \sum_{i=1}^{I} \left\{ \frac{1}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} \right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma(1-\eta)}{\gamma-\gamma\theta+\theta}} \right\}^{\frac{\theta-1}{\eta-1}} \right\}^{\frac{\gamma-\gamma\theta+\theta}{\theta-1}}$$

Because the mapping is monotone (under some conditions), this equation has a unique solution $W_Y^\star$.

2. Using s_{ij}, μ_{ij} and $W^\gamma Y^{1-\gamma}$, compute firm-level prices p_{ij} from the pricing kernel:

$$p_{ij} = W_Y^{\frac{1}{\gamma-\gamma\theta+\theta}} \left\{ \frac{1}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} P^\theta\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma}{\gamma-\gamma\theta+\theta}}$$

or directly from:

$$p_{ij} = \frac{\left\{ \frac{1}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}} P^\theta\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma}{\gamma-\gamma\theta+\theta}}}{\left\{ \frac{1}{J} \int \left\{ \frac{1}{I} \sum_{i=1}^{I} \left\{ \frac{1}{\gamma} \frac{\left[\frac{1}{J} I^{\frac{1-\theta}{\eta-1}} s_{ij}^{\frac{\eta-\theta}{\eta-1}}\right]^{\frac{1}{\gamma}-1}}{A_{ij}^{1/\gamma} \mu_{ij}^{-1}} \right\}^{\frac{\gamma(1-\eta)}{\gamma-\gamma\theta+\theta}} \right\}^{\frac{\theta-1}{\eta-1}} \right\}^{\frac{1}{1-\theta}}}$$

(iii) **Solve the aggregate block.**

1. Use W_Y and solve the labor-market clearing condition:

$$L^d(W, Y) = L^s(W),$$

where

$$Y^{\frac{1}{\gamma}} J^{-\frac{1}{\gamma}} \int_j I^{\frac{\theta-1}{\gamma(1-\eta)}} \sum_i \left(\frac{s_{ij}^{\frac{\theta-\eta}{1-\eta}} p_{ij}^{-\theta}}{A_{ij}} \right)^{\frac{1}{\gamma}} dj + \int_j \sum_i \phi_{ij} = \bar{\varphi} \left(\frac{W}{P} \right)^{\varphi}.$$

Solve this monotone equation numerically for Y.

2. Get Y from $Y = W_Y / W^\gamma$.

3. Get L from $L^d(W, Y)$ using $\{W, Y\}$

(iv) **Back out firm-level variables.**

1. With $\{W, Y, L\}$ determined, recover $\{y_{ij}, l_{ij}\}$ using the demand function $y_{ij} = \frac{1}{J} \frac{1}{N_j} p_{ij}^{-\eta} p_j^{\eta-\theta} P^\theta Y$ and the production function $l_{ij} = \left(\frac{y_{ij}}{A_{ij}}\right)^{\frac{1}{\gamma}}$.

B Additional Figures

B.1 Measures of Market Power

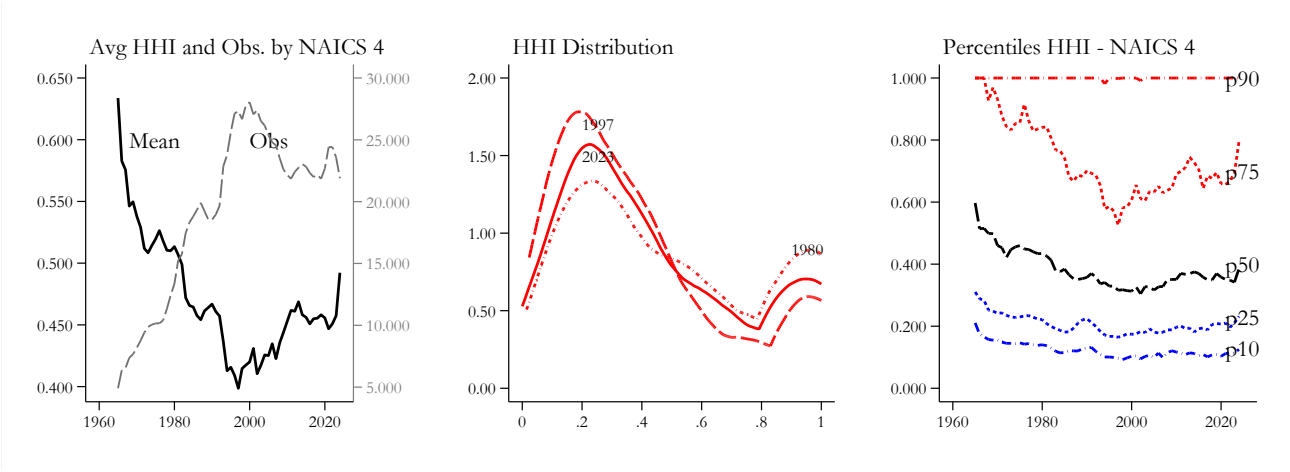


Figure B.1: HHI Over Time and Distribution

Note: In the left panel, the left axis refers to the average HHI across industries at the NAICS4 level, while the right axis refers to the number of observations. The center panel shows the distribution of HHI in 1980, 2023, and the minimum HHI year (1997). In the right panel, blue indicates below-median percentiles and red indicates above-median percentiles.

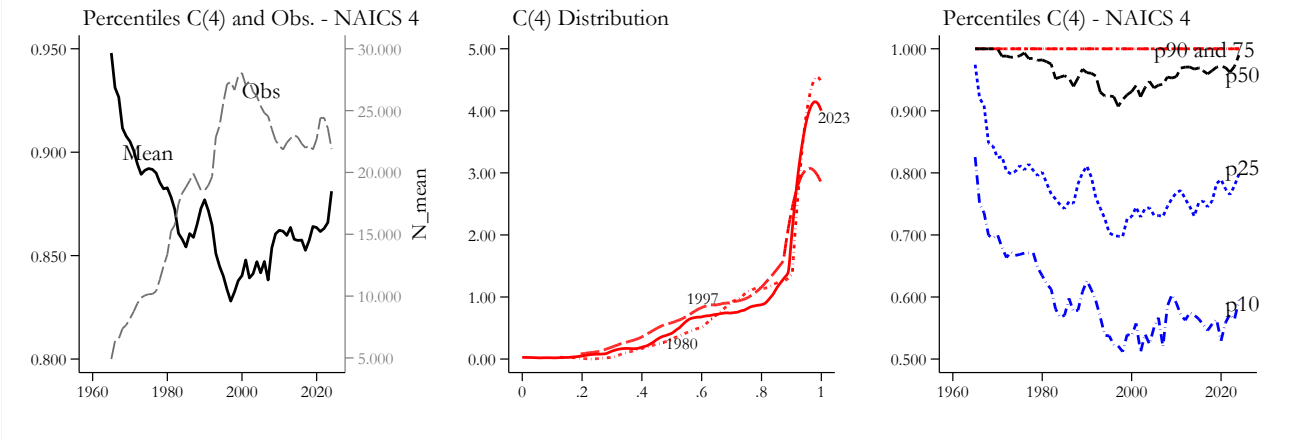


Figure B.2: C(4) over time and distribution

Note: In the left panel, the left axis refers to the average C(4) across industries at the NAICS4 level, while the right axis refers to the number of observations. The center panel shows the distribution of C(4) in 1980, 2023, and the minimum C(4) year (1997). In the right panel, blue indicates below-median percentiles and red indicates above-median percentiles.

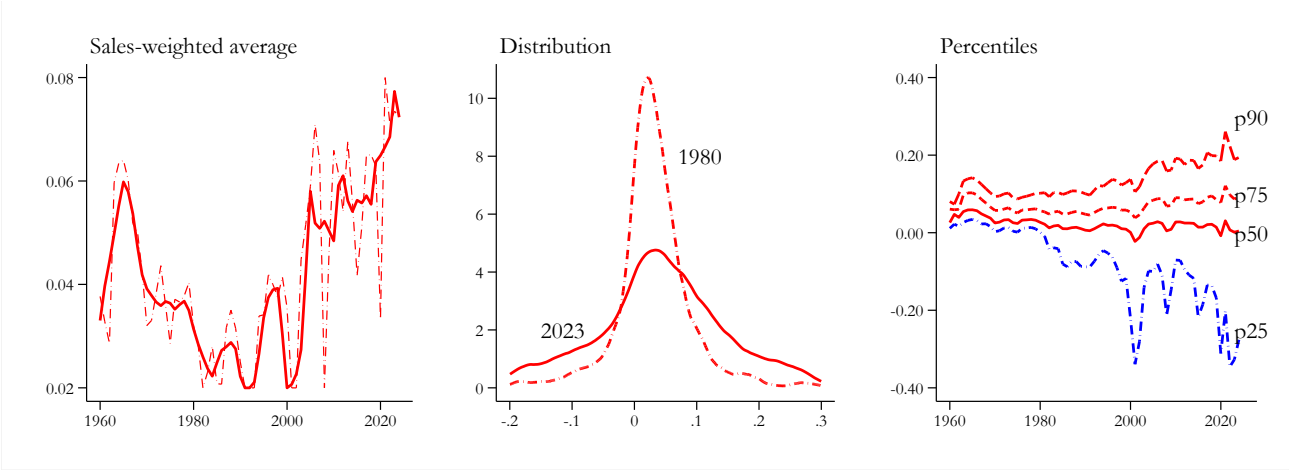


Figure B.3: Accounting Profit Rate $\frac{\Pi_{it}}{R_{it}}$

Note: The solid line in the left panel is the 5 year moving average (5YMA). The center panel limits the range of the plot of the Profit Rate within [-0.2, 0.3]. In the right panel, blue indicates below-median percentiles and red indicates above-median percentiles.

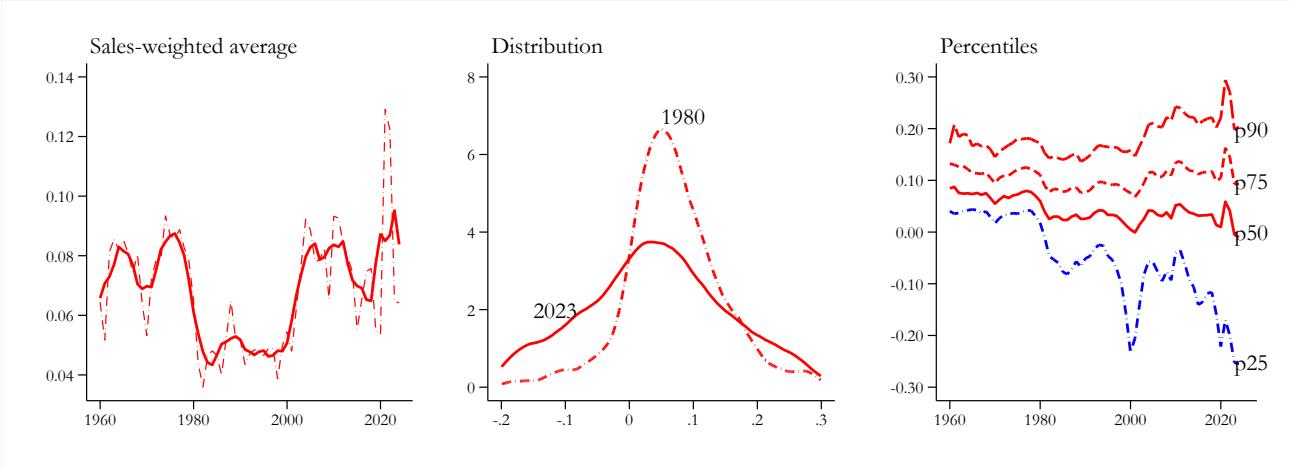


Figure B.4: Imputed Profit Rate $\frac{\Pi_{it}}{R_{it}}$

Note: The solid line in the left panel is the 5 year moving average (5YMA). The center panel limits the range of the plot of the Profit Rate within [-0.2, 0.3]. In the right panel, blue indicates below-median percentiles and red indicates above-median percentiles.

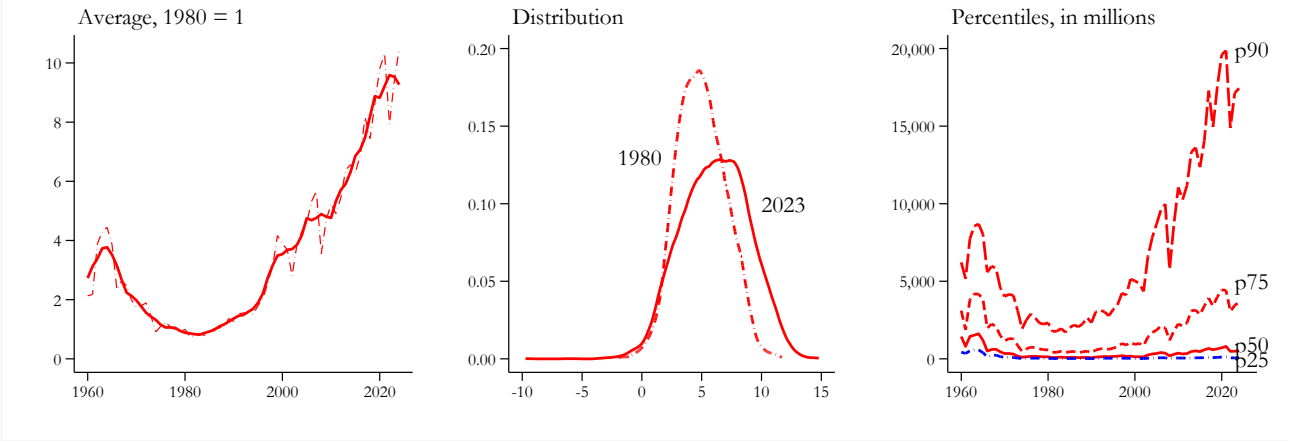


Figure B.5: Market Value over time and distribution

Note: The center panel limits the range of the plot of the market value within [-10, 15]. In the right panel, blue indicates below-median percentiles and red indicates above-median percentiles.

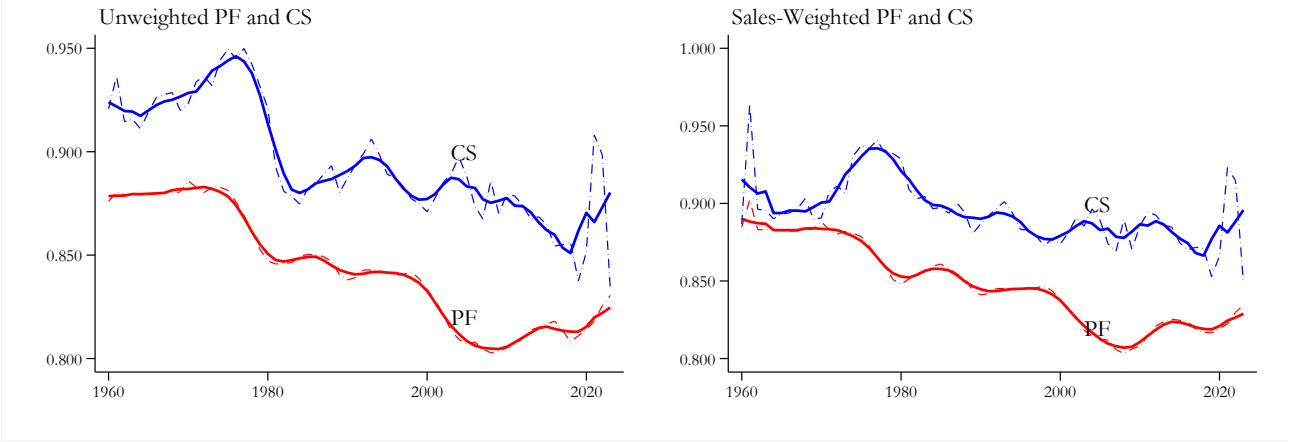


Figure B.6: Production Approach and Cost Share - Unweighted vs Sales-weighted

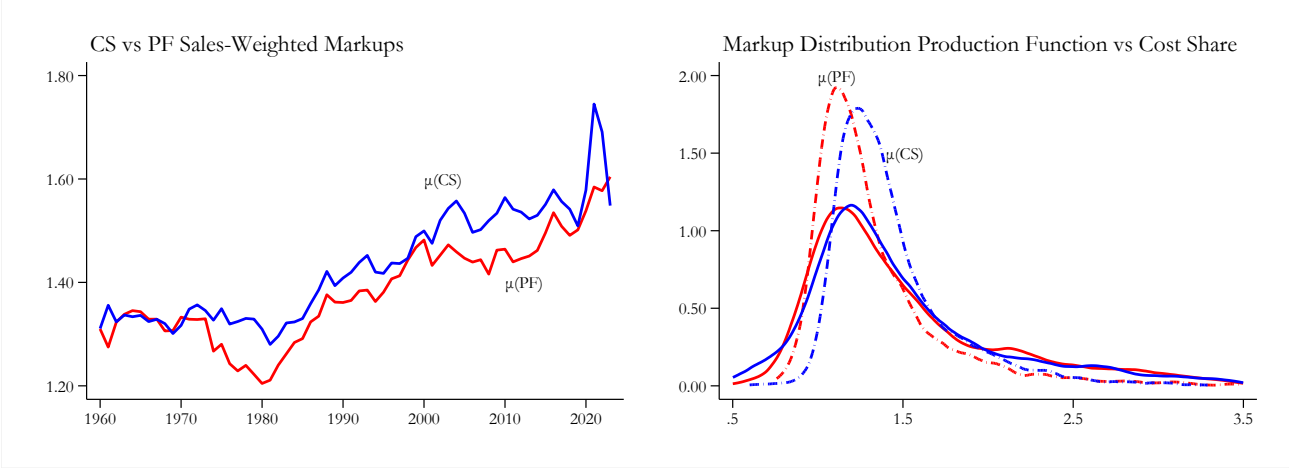


Figure B.7: Markup with Cost Share and Production Approach: SW-average and distribution

Note: The right panel limits the range of the plot of the markup within [0.5, 3.5]; Blue lines represent Cost Share markups; red lines represent Production Approach markups. Solid lines refer to 2023, dashed lines to 1980. The costs are based on the imputed user cost of capital.

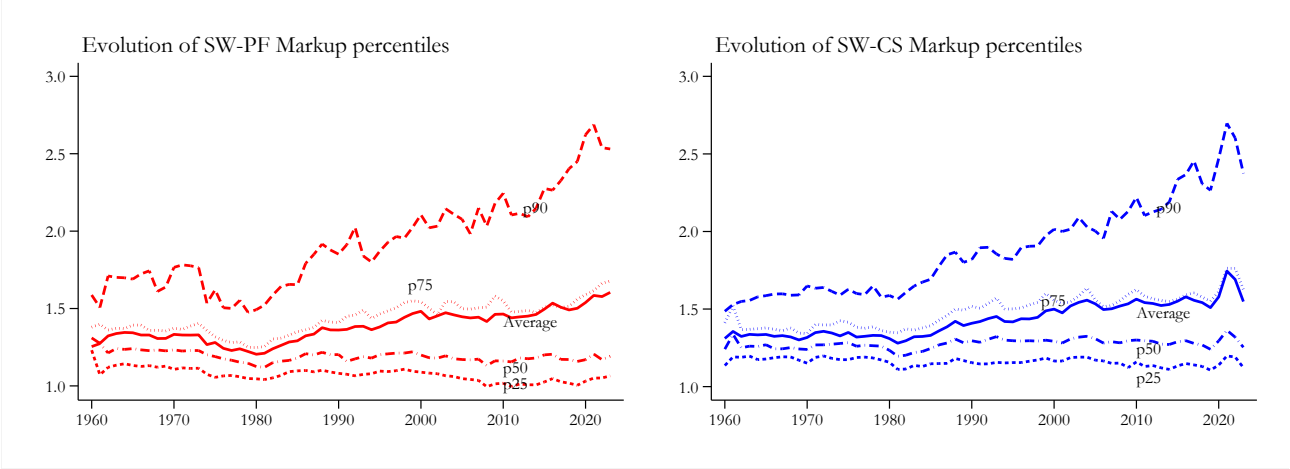


Figure B.8: Percentile trends of sales-weighted markups over time - Production Approach and Cost Shares.

Note: Blue lines represent Cost Share markups; red lines represent Production Approach markups. Cost based on imputed user cost of capital.

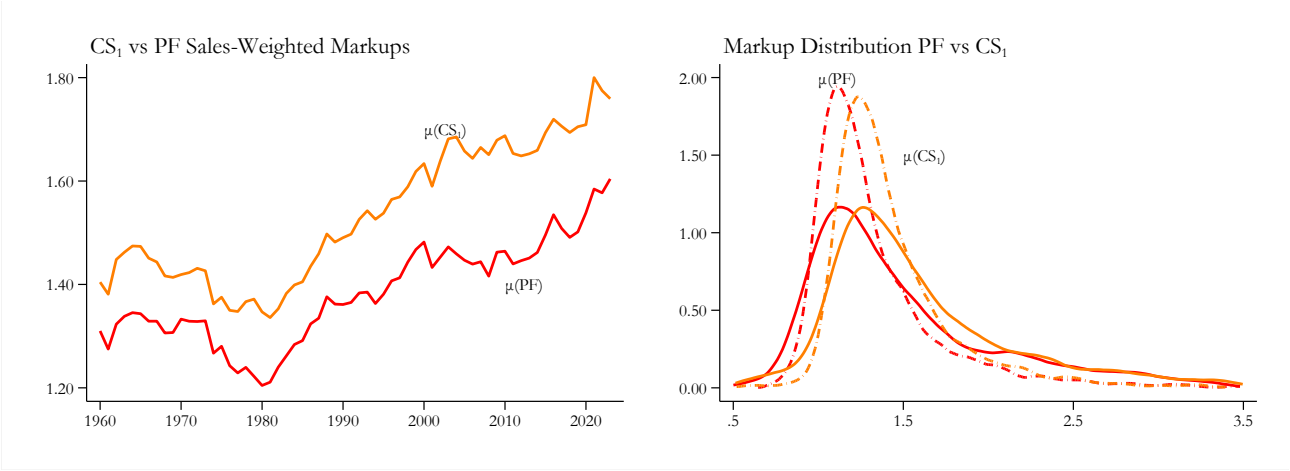


Figure B.9: Markups with Costs Share (new k_{exp}) versus Production Approach: SW-average and distribution

Note: The right panel limits the range of the plot of the markup within [0.5, 3.5]; Orange lines represent Cost Share markups; red lines represent Production Approach markups. Solid lines refer to 2023, dashed lines to 1980. Capital Expenditure has been defined as $k_{exp} = at\frac{xint}{lt} = at \cdot \text{rate}$.

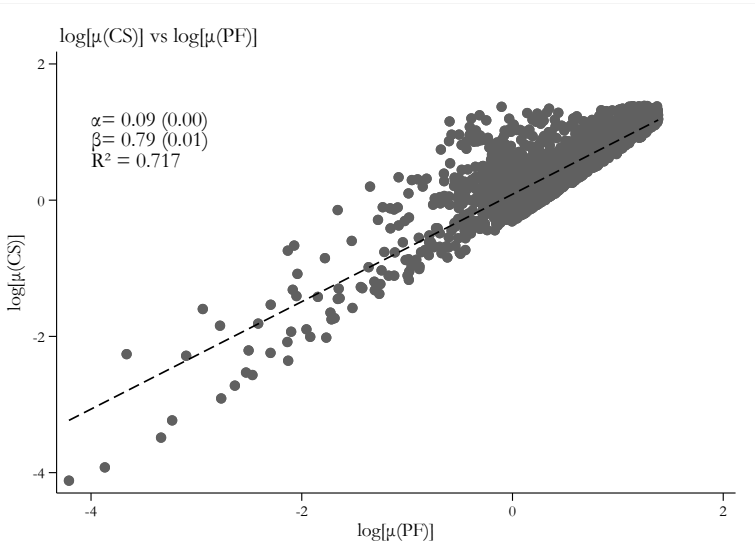


Figure B.10: Log-Log scatter plot and linear fit of Production Approach vs. Cost Share markups in 2023.

Note: The dashed line indicates the linear regression fit. We delimit the markup between [0,4].

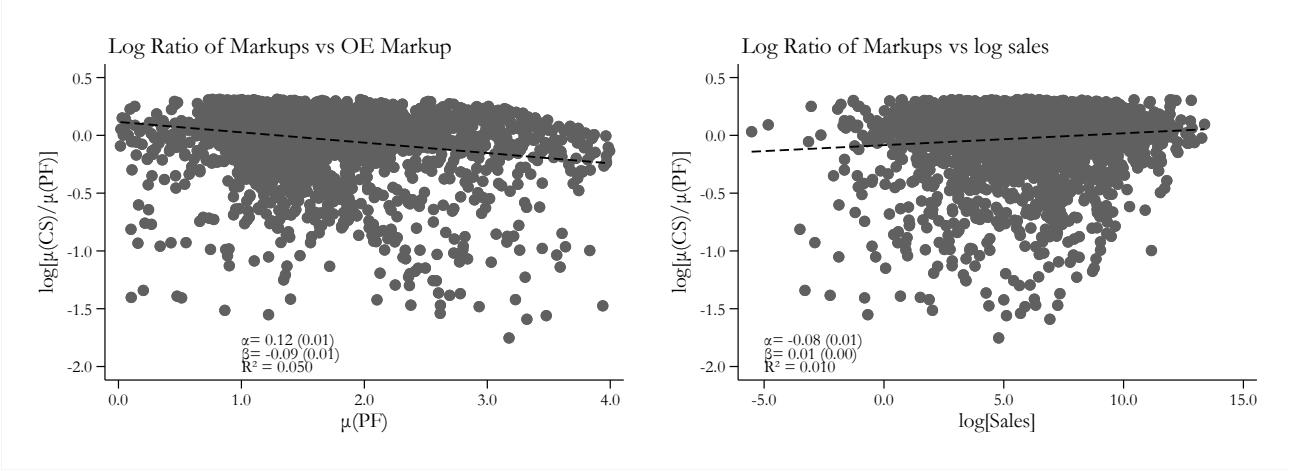


Figure B.11: Cross-sectional comparison of markup differences in 2023.

B.2 Comparative Statics

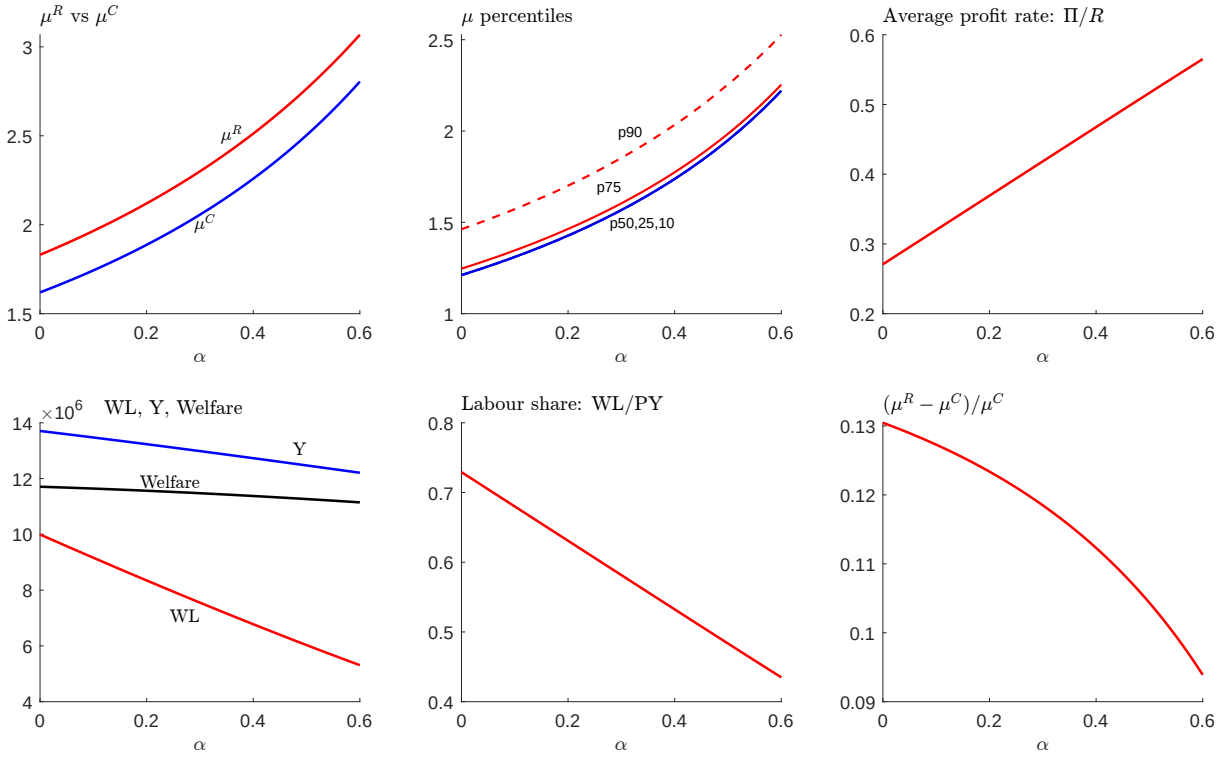


Figure B.12: Comparative Statics in Common Ownership share $\alpha \in [0, 0.6]$

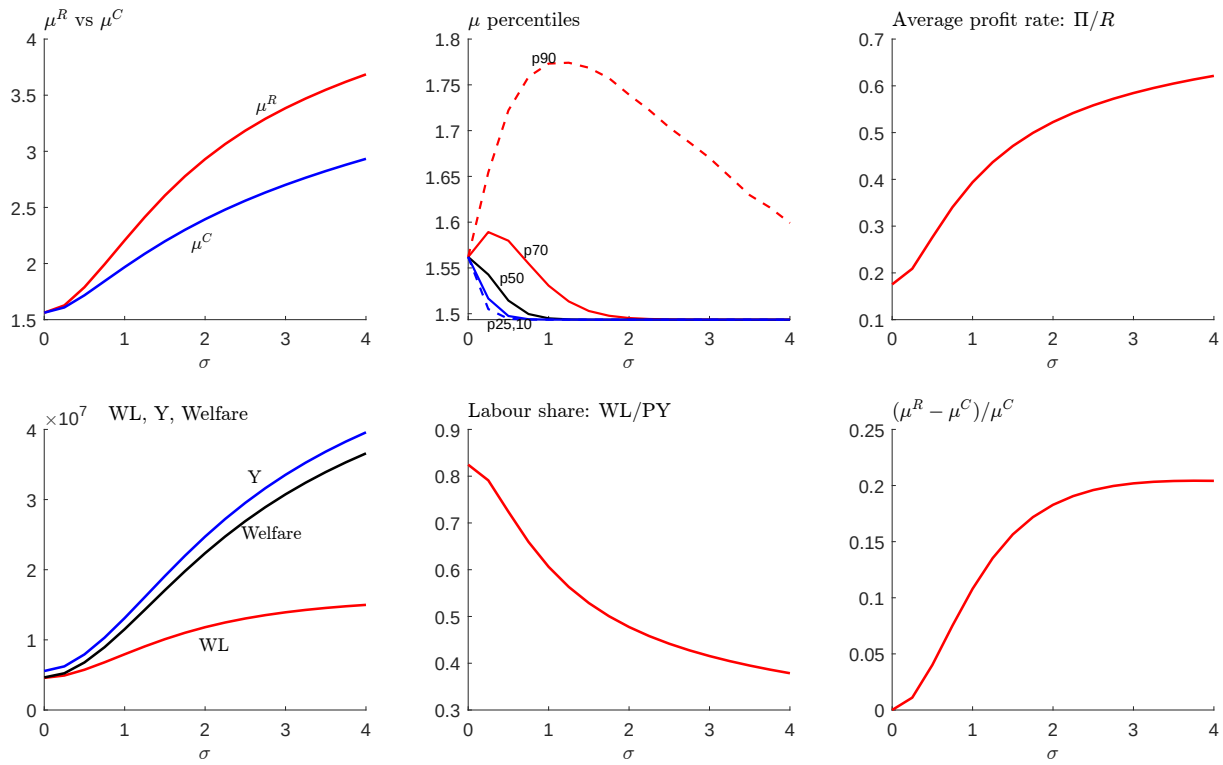


Figure B.13: Standard deviation of Firm-level (log) TFP shock share $\sigma \in [0, 4]$

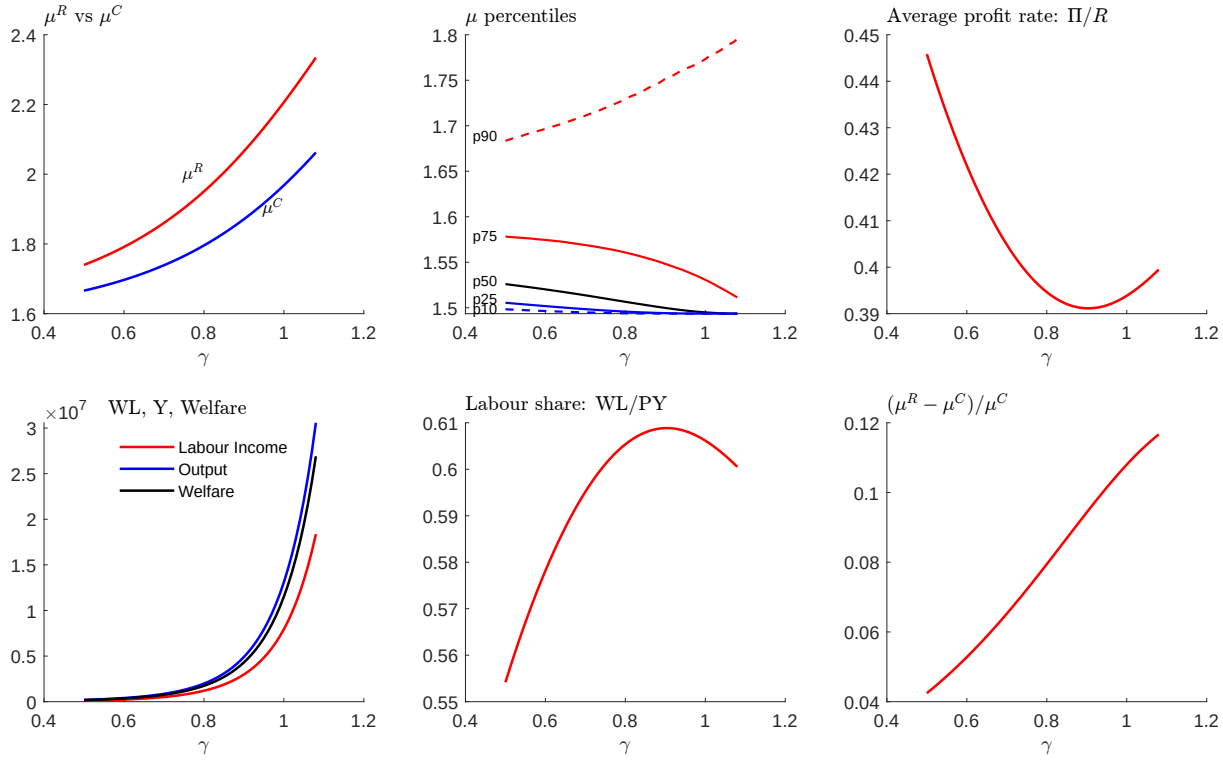


Figure B.14: Return to scale $\gamma \in [0, 1, 1.05]$

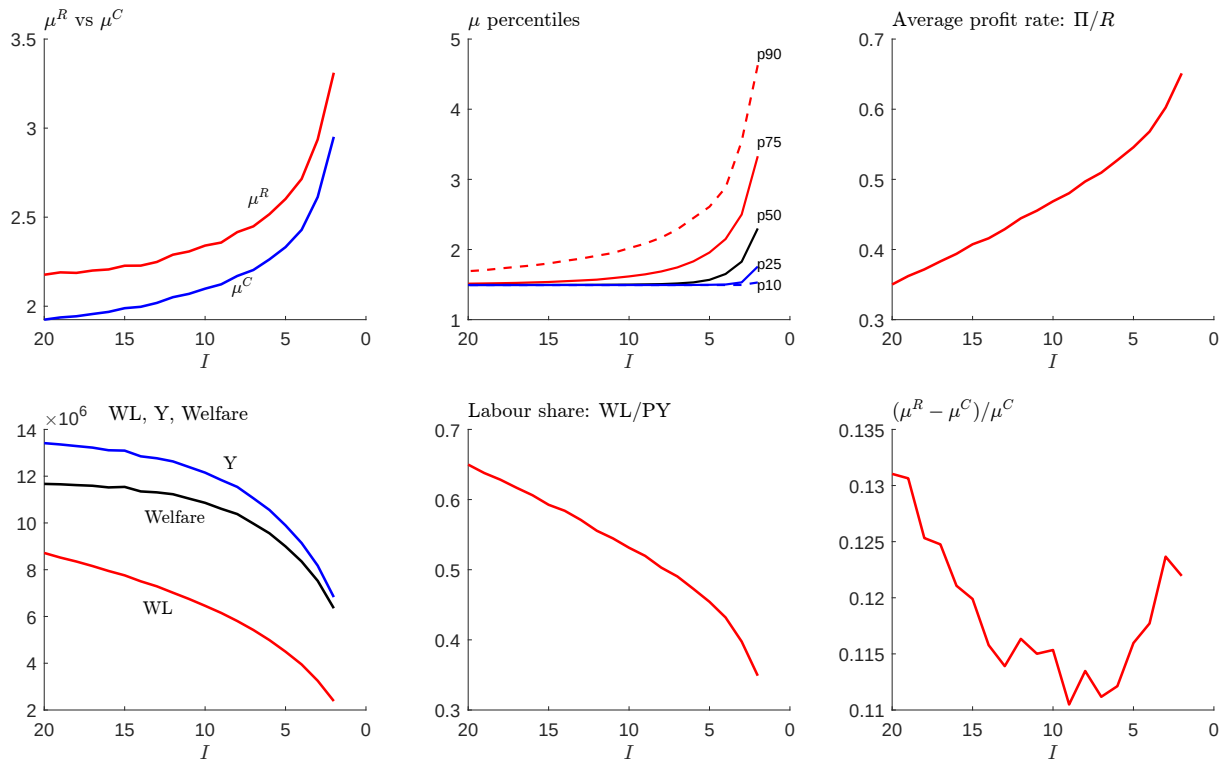


Figure B.15: Number of firms $I \in \{2, 3, ..., 20\}$

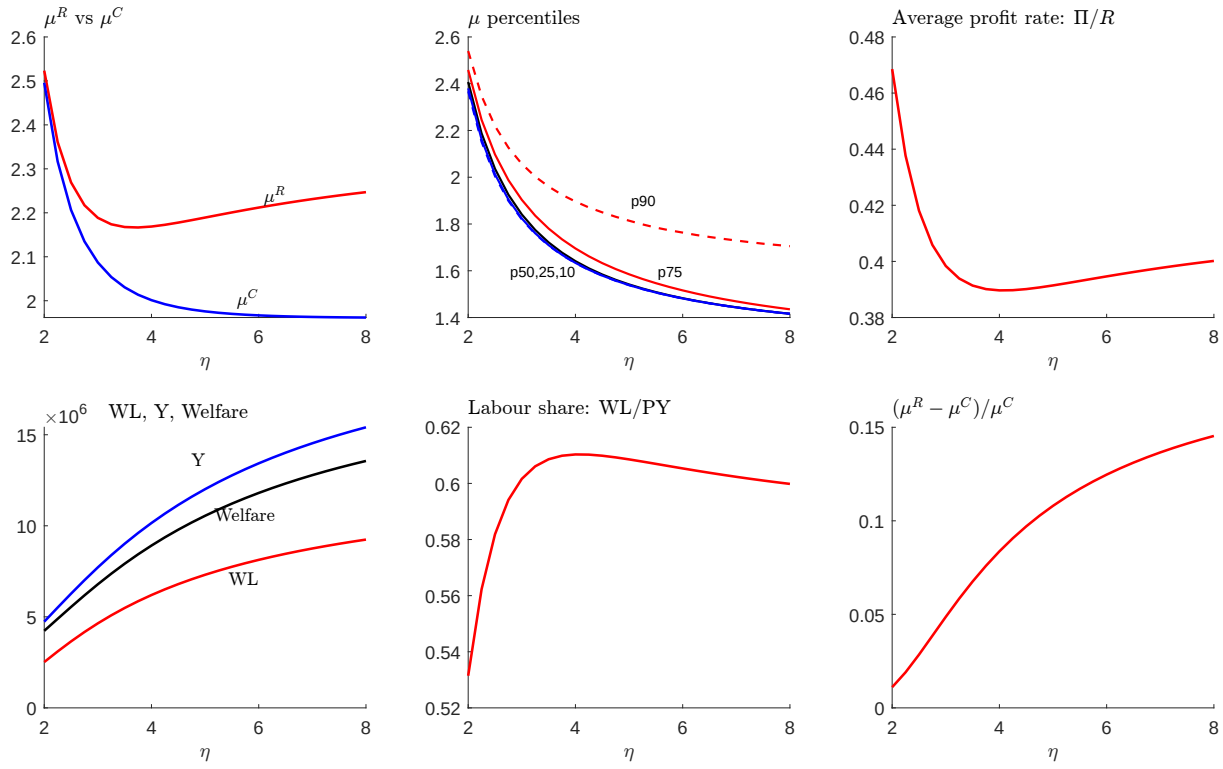


Figure B.16: Within-market elasticity $\eta \in [2, 8]$

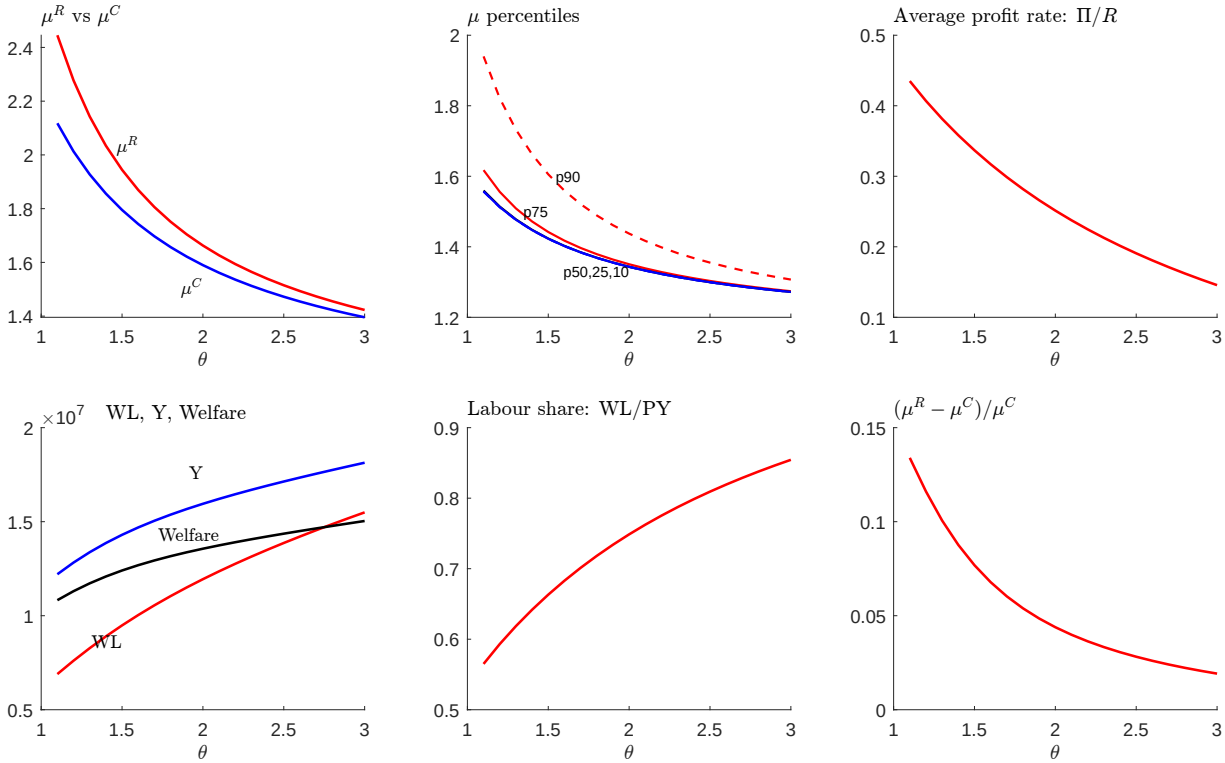


Figure B.17: Between-market elasticity $\theta \in [1.1, 3]$

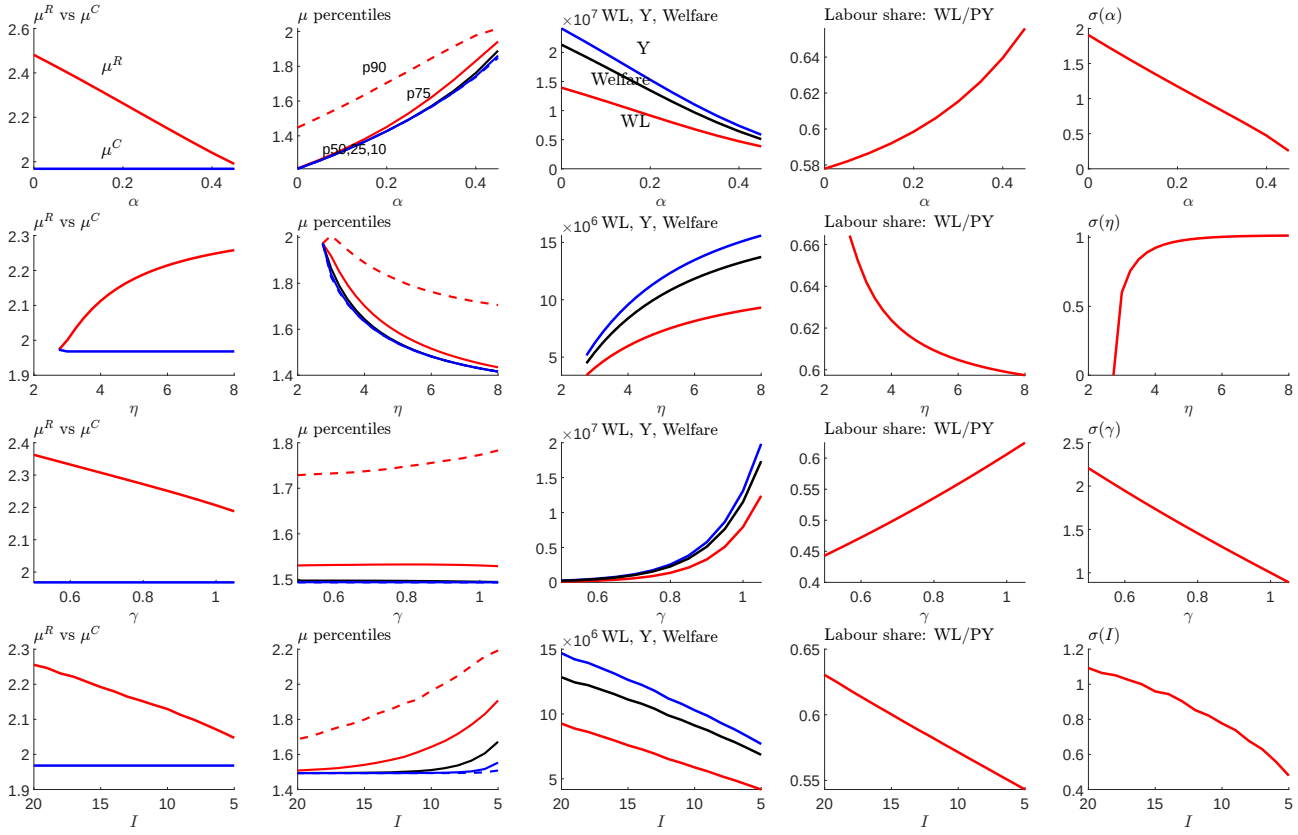


Figure B.18: Comparative statics to maintain a constant cost-weighted markup.

Notes. We first compute the cost-weighted markup for the benchmark case defined by the parameter set in Table 1, then vary one parameter at a time and solve for the value of σ that keeps the cost-weighted markup constant. Finally, we plot the resulting comparative statics.

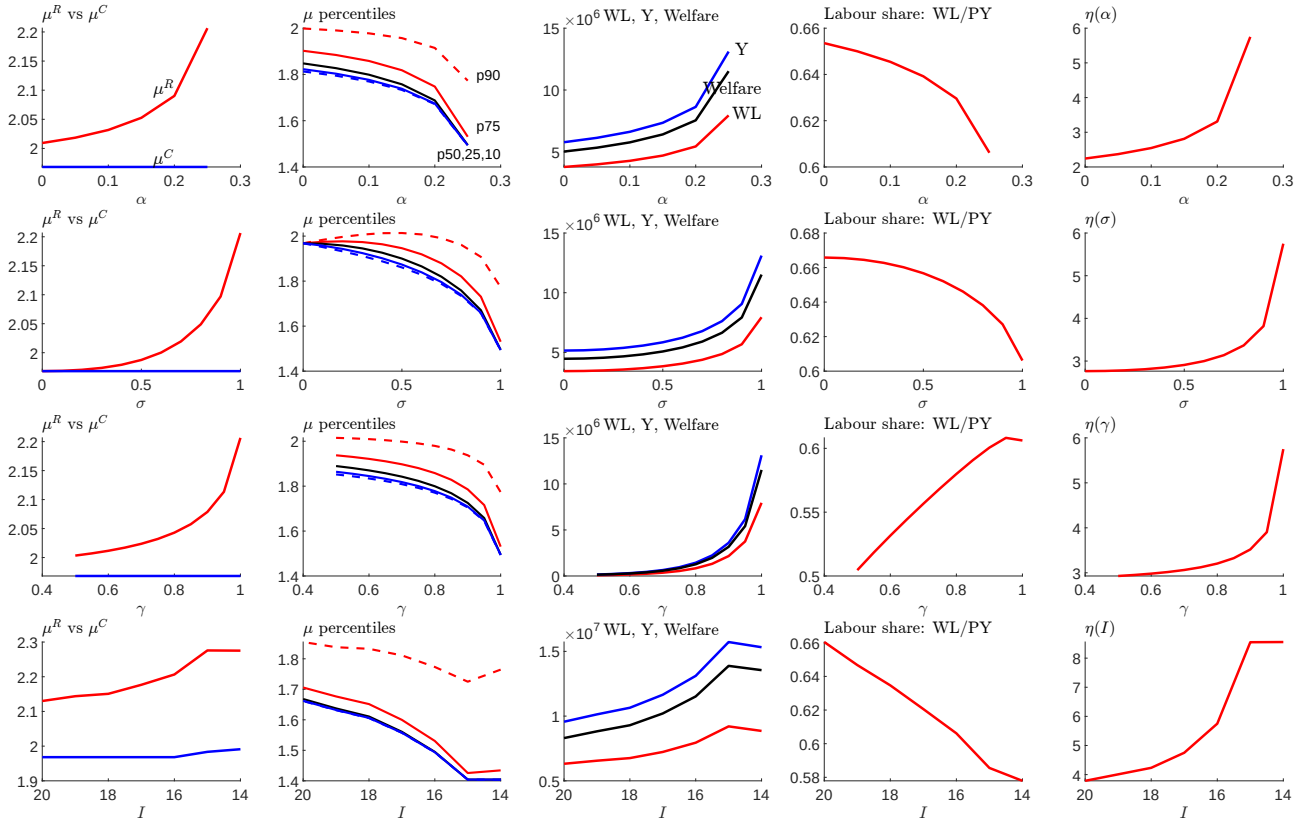


Figure B.19: Comparative statics to maintain a constant cost-weighted markup.

Notes. We first compute the cost-weighted markup for the benchmark case defined by the parameter set in Table 1, then vary one parameter at a time and solve for the value of η that keeps the cost-weighted markup constant. Finally, we plot the resulting comparative statics.

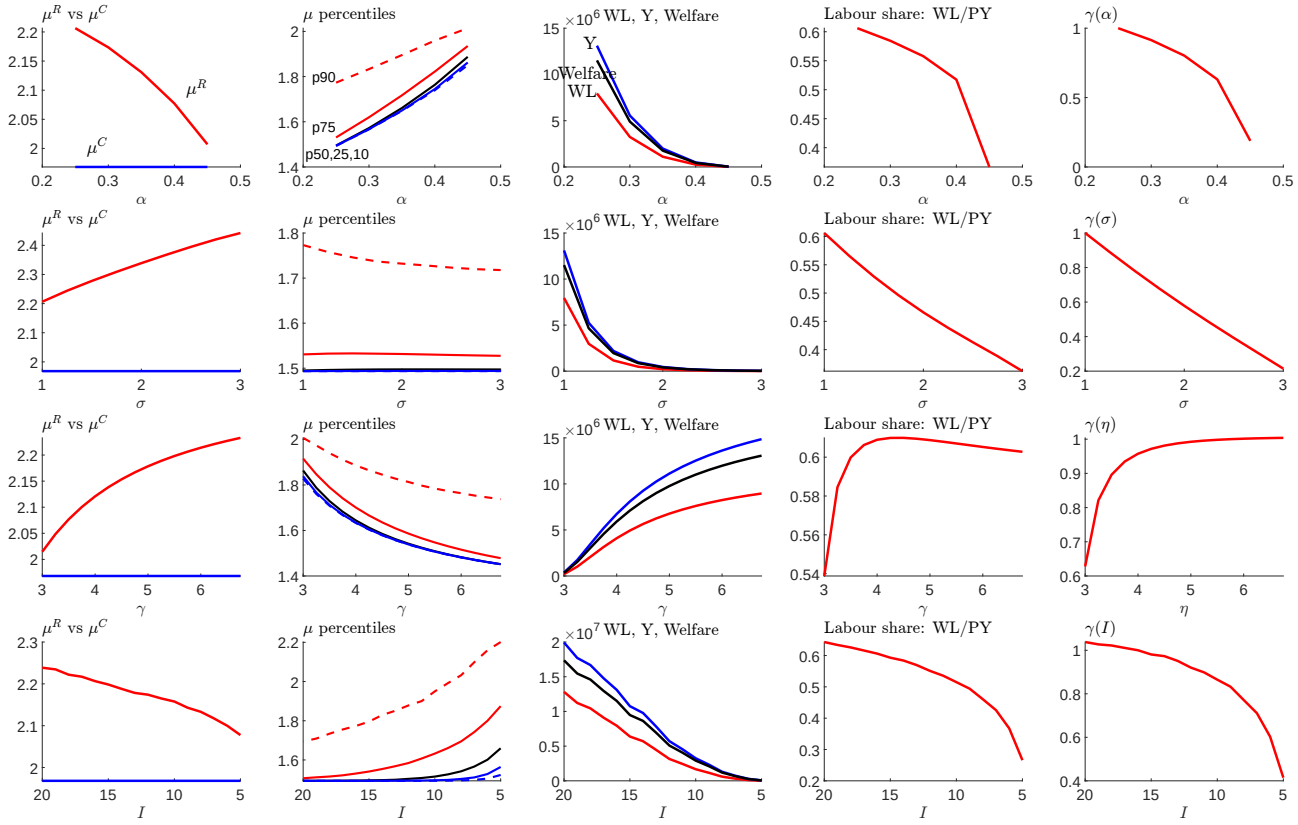


Figure B.20: Comparative statics to maintain a constant cost-weighted markup.

Notes. We first compute the cost-weighted markup for the benchmark case defined by the parameter set in Table 1, then vary one parameter at a time and solve for the value of γ that keeps the cost-weighted markup constant. Finally, we plot the resulting comparative statics.

C Markup Decomposition

We decompose the markup in equation (4). We introduce some additional notation and define:

$$\Delta\mu_t \equiv \mu_t - \mu_{t-1}, \quad \Delta m_{it} \equiv m_{it} - m_{i,t-1}, \tag{C.7}$$

$$\Delta\theta_{it} \equiv \theta_{it} - \theta_{i,t-1}, \quad \Delta\kappa_{it} \equiv \kappa_{it} - \kappa_{i,t-1}, \tag{C.8}$$

$$\tilde{m}_{it} = m_{it} - m_{t-1}, \quad \tilde{m}_{j,t-1} = m_{j,t-1} - m_{t-1}. \tag{C.9}$$

$$\theta_{t-1} = \sum_{i \in N_{t-1}} m_{i,t-1}\theta_{i,t-1}, \quad \kappa_{t-1} = \sum_{i \in N_{t-1}} m_{i,t-1}\kappa_{i,t-1}, \quad m_{t-1} = \frac{1}{N_{t-1}}. \tag{C.10}$$

$$
\begin{aligned}
\Delta\mu_t = & \underbrace{\underbrace{\sum_{i\in C} m_{i,t-1}\theta_{i,t-1}\Delta\kappa_{it}}_{\text{Continuation firms}} + \underbrace{\left[\sum_{i\in \text{Entry}} m_{it}\theta_{t-1}\tilde{\kappa}_{it} - \sum_{j\in \text{Exit}} m_{j,t-1}\theta_{t-1}\tilde{\kappa}_{j,t-1}\right]}_{\text{Net entry}}}_{\text{Within-}\kappa} \\
& + \underbrace{\underbrace{\sum_{i\in C} m_{i,t-1}\kappa_{i,t-1}\Delta\theta_{it}}_{\text{Continuation firms}} + \underbrace{\left[\sum_{i\in \text{Entry}} m_{it}\kappa_{t-1}\tilde{\theta}_{it} - \sum_{j\in \text{Exit}} m_{j,t-1}\kappa_{t-1}\tilde{\theta}_{j,t-1}\right]}_{\text{Net entry}}}_{\text{Within-}\theta} \\
& + \underbrace{\underbrace{\sum_{i\in C} \theta_{i,t-1}\kappa_{i,t-1}\Delta m_{it}}_{\text{Continuation firms}} + \underbrace{\left(\sum_{i\in \text{Entry}} m_{it} - \sum_{j\in \text{Exit}} m_{j,t-1}\right)\theta_{t-1}\kappa_{t-1}}_{\text{Net entry change in market share}}}_{\text{Reallocation}} \\
& + \underbrace{\underbrace{\sum_{i\in C} m_{i,t-1}\Delta\theta_{it}\Delta\kappa_{it}}_{\text{Continuation firms}} + \underbrace{\left[\sum_{i\in \text{Entry}} m_{it}\tilde{\theta}_{it}\tilde{\kappa}_{it} - \sum_{j\in \text{Exit}} m_{j,t-1}\tilde{\theta}_{j,t-1}\tilde{\kappa}_{j,t-1}\right]}_{\text{Net entry}}}_{\text{Cross}(\kappa,\theta)} \\
& + \underbrace{\sum_{i\in C} \kappa_{i,t-1}\Delta m_{it}\Delta\theta_{it}}_{\text{Cross}(m,\theta)} + \underbrace{\sum_{i\in C} \theta_{i,t-1}\Delta m_{it}\Delta\kappa_{it}}_{\text{Cross}(m,\kappa)} + \underbrace{\sum_{i\in C} \Delta m_{it}\Delta\theta_{it}\Delta\kappa_{it}}_{\text{Triple interaction}}.
\end{aligned}
\tag{C.11}
$$