Price Markups or Wage Markdowns?[*]

Kevin Ren[†] Dalton Rongxuan Zhang[‡]

First Version: January 31, 2025
This Version: May 13, 2025

Abstract

This paper revisits the rise in aggregate market power by decomposing it into product and labor market components. We address three key questions empirically and quantitatively: (i) Has the increase in market power been primarily driven by price markups? (ii) Do larger firms exhibit higher price markups? (iii) How does incomplete price pass-through relate to firms' input market power? Using U.S. Census data, we find that the aggregate price markup has remained stable, while labor market power, measured by the aggregate wage markdown, has nearly doubled. Moreover, price markups are negatively related to firm size, contradicting the conventional assumption of Marshall's Second Law of Demand, whereas wage markdowns are positively related. To explain these findings, we develop a general equilibrium model of monopolistic-monopsonistic competition with nonparametric product demand and labor supply. The model shows that incomplete price pass-through, often interpreted as evidence of higher markups among larger firms, can instead result from monopsony power. Firms with labor market power lower wages in response to negative shocks, dampening cost increases and generating incomplete pass-through. Our results suggest that labor market power may be the primary force behind the observed patterns of incomplete pass-through, the negative labor share-market power relationship, and the broader rise in aggregate market power.

[*]We would like to thank our advisors George-Marios Angeletos, Lawrence Christiano, Nicolas Crouzet, Martin Eichenbaum, Diego Känzig, Dimitris Papanikolaou, Matthew Rognlie, and Bryan Seegmiller for their guidance and support. We would also like to thank Efraim Benmelech, Vivek Bhattacharya, Joel David, Maarten De Ridder, Erik Hurst, Gaston Illanes, Evan Majic, Kiminori Matsuyama, Konstantin Milbradt, Cristoforo Pizzimenti, Jonathan Rothbaum, Kunal Sangani, Lawrence Schmidt, Aditya Soenarjo, and Chad Syverson for helpful discussions and comments. The U.S. Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7503840 (P-7518691), Disclosure Review Board (DRB) approval numbers: CBDRB-FY25-SEHSD003-016, CBDRB-FY25-0271).

[†]Kellogg School of Management, Northwestern University. kevin.ren@kellogg.northwestern.edu

[‡]Department of Economics, Northwestern University. r.zhang@u.northwestern.edu

1 Introduction

There is significant evidence of a secular increase in aggregate market power over the last few decades in the United States, which is usually attributed to increases in product market power. Cross-sectionally, a crucial relationship for both welfare and policy analysis lies between firm size and markups.[1] The conventional assumption, often referred to as Marshall's Second Law of Demand, posits that larger firms have higher markups. This notion is seemingly supported by the observed phenomenon of incomplete price pass-through. In this paper, we revisit these two foundational observations by decomposing market power into its product and labor market components. This approach leads us to three core questions: Has the increase in market power been primarily driven by price markups? Do larger firms exhibit higher price markups? How does incomplete price pass-through relate to firms' input market power?

We address these questions both empirically and theoretically. Empirically, we revisit the question posed by De Loecker, Eeckhout and Unger (2020) by studying the systematic patterns of market power across firms and over time. We extend their analysis by allowing for the possibility that firms exert labor market power in addition to product market power. Using the U.S. Census Bureau's Longitudinal Business Database (LBD), which contains establishment-level employment and wage data for the universe of U.S. non-farm private sector firms, we disentangle labor market power from what has previously been implicitly considered "total market power" at the firm level. A key distinction of our approach is that we exclude the wage bill from the set of "flexible" inputs that were previously assumed to be competitively priced, thereby allowing for the presence of labor market power.

After accounting for labor market power, we find that the aggregate price markup has remained stable over time. The aggregate wage markdown, however, has increased substantially, nearly doubling over the sample period. Decomposing changes in aggregate labor market power into changes in the estimated aggregate labor output elasticity and directly measured the aggregate cost share, we find that labor output elasticity has decreased by roughly 23% at its peak, consistent with secular changes in production factors, such as automation.[2] However, the labor cost share itself has declined by as much as 60% percent, indicating a substantial rise in labor market power despite the drop in labor output elasticity.[3]

The rise in the aggregate wage markdown mirrors the findings of De Loecker, Eeckhout and Unger (2020) for markups and appears to be driven by firms at the top of the distribution. Specifically, both the standard deviation of markdowns and the gap between the 90th and 50th percentiles have increased markedly, whereas the 50th-10th percentile gap has grown only modestly.

[1] We use the terms *markups* and *markdowns* interchangeably with price markups and wage/input markdowns, respectively, throughout the paper.

[2] Labor cost share is defined as the total wage bill divided by total revenue at the firm level.

[3] The large decline in the directly measured labor cost share suggests our conclusion is likely robust to alternative specifications or potential misspecifications of the production function; Section 4.2 discusses this in greater detail.

Markups follow the same pattern, albeit on a much smaller scale. Together with the cross-sectional evidence presented in Section 4, these patterns point to a key role for reallocation: larger firms, which have higher markdowns, have expanded their market share, and inputs such as labor have shifted increasingly toward them. To probe this channel further, we examine productivity dispersion over our sample. From the first to the last year, the standard deviation of log value-added per worker has doubled, and the difference between the 90th and 50th percentiles of log labor productivity has nearly tripled. In other words, the most productive firms are becoming even more productive and dominant. Finally, we show in a calibrated model how the trends of increased markdown dispersion, reallocation toward high-markup firms, and rising productivity dispersion are interconnected.

Turning to the cross-sectional evidence, we find that after decomposing market power into its product and labor components, markups are *negatively* related to firm size measures such as labor productivity, product and labor market shares, and others. In contrast, markdowns are *positively* related to these measures. However, total market power, which combines both labor and product market power, is *positively* related to firm size measures. This suggests that the conventional assumption regarding Marshall's Second Law of Demand, which states that price elasticity decreasing with quantity demanded or that larger firms have higher markups may not hold in the data. Moreover, the associated empirical evidence supporting Marshall's Second Law of Demand may actually be driven by the existence of labor market power.

To reconcile our finding regarding Marshall's Second Law of Demand, we develop a flexible general equilibrium model with monopolistic-monopsonistic competition, incorporating nonparametric product demand and labor supply systems. This framework allows us to analyze the competitive environment in both product and labor markets. Traditionally, conclusions about Marshall's Second Law of Demand have been based on measured incomplete price pass-through: if firms do not fully pass cost shocks onto prices, their markups shrink in response to negative TFP or cost shocks, implying that larger firms have higher markups. Our cross-sectional evidence suggests a different interpretation of the link between incomplete price pass-through and Marshall's Second Law of Demand. Specifically, we show that an upward-sloping labor supply curve dampens the transmission of TFP and cost shocks into prices. This mechanism can generate incomplete price pass-throughs as measured in the data without Marshall's Second Law of Demand holding. The intuition is as follows: a firm with labor market power optimally chooses a point along the labor supply curve rather than taking wages as given. When a negative TFP shock hits, the firm raises its price and demand falls. With lower demand, the firm needs fewer workers and hence lower wages. The resulting wage reduction lowers marginal costs, which in turn limits the rise in prices. The quantitative importance of this channel depends on the demand and labor supply system, and we explore its magnitude in detail using our calibrated framework.

We use our model to nonparametrically identify the product demand and labor supply systems based on data from 1977. Consistent with our empirical findings, the identified demand system exhibits what Matsuyama (2025) describes as Marshall's *anti*-Second Law of Demand: demand elasticity increases with quantity, and price pass-through exceeds one with respect to changes in *marginal cost*.[4] In contrast, the labor supply system follows a more conventional pattern found in the labor economics literature: supply elasticity decreases with the quantity of supply, or larger firms have larger wage markdowns. When we analyze the demand system-implied price pass-through, we find that effective price pass-throughs are significantly below 1, and that larger, more productive firms exhibit lower effective price pass-throughs due to their inelastic labor supply. This finding aligns with empirical evidence such as Amiti, Itskhoki and Konings (2019).

Building on these insights, we next explore whether rising productivity dispersion and the emergence of "superstar firms" help explain the increase in the aggregate wage markdown. Motivated by our finding that the aggregate wage markdown is driven primarily by firms in the upper percentiles and by the substantial rise in productivity dispersion we feed each year's observed firm productivity distribution into the model while holding the underlying demand and supply systems fixed. Relative to 1977, which is the start of the sample, there are both more superstar firms and larger productivity gaps between them and the remaining firms. We find that the rise in TFP dispersion accounts for roughly 30% of the observed increase in the aggregate markdown, while the aggregate markup remains stable or declines slightly. Digging deeper, revenue shares per firm have fallen for both high- and low-productivity firms, and individual changes in markup and markdown levels are small. However, because a greater share of economic activity is concentrated in high-productivity firms, the aggregate markdown rises. These results indicate that reallocation toward superstar firms is the primary driver of the trends we document.

Related Literature. This paper contributes to the growing body of literature on aggregate market power by simultaneously addressing its product and labor market dimensions. A large strand of research has focused on measuring firm-level markups or product market power and understanding their aggregate implications (e.g. De Loecker and Warzynski, 2012; Autor et al., 2017; Bustamante and Donangelo, 2017; Gutiérrez and Philippon, 2017; Corhay, Kung and Schmid, 2020; De Loecker, Eeckhout and Unger, 2020; Autor et al., 2020; Crouzet and Eberly, 2021, 2023; De Ridder, 2024). These studies often emphasize the rise in aggregate markups over recent decades.

Our analysis contributes to a growing body of research that empirically estimate labor market power while purging out product market power (e.g. Dobbelaere and Mairesse, 2013; Morlacco, 2020; Traina, 2018; Yeh, Macaluso and Hershbein, 2022; Rubens, 2023). In particular, our goal of studying wage markdowns at a macroeconomic level and thus recovering these measures across a broad set of industries makes our work most comparable to Yeh, Macaluso and Hershbein (2022) and Traina (2018), which both estimate wage markdowns for U.S. manufacturing plants. Our approach, however, differs in two key respects. First, our dataset covers publicly traded firms across almost all U.S. private sectors (excluding finance and utilities) rather than solely focusing on

[4] Note that wages and markdowns are also components of marginal cost.

the manufacturing sector, permitting a more comprehensive macro-level perspective. Second, by employing firm-level data, we capture high-skill workers who are not employed on plant floors but rather in settings such as corporate headquarters or other non-plant settings. Empirically, Yeh, Macaluso and Hershbein (2022) and Traina (2018) find that markdowns for manufacturing plants rose by about 10% and 100% respectively. Our results provide evidence for other sectors of the U.S. economy and shows that the aggregate markdown increase closer to the estimate of Traina (2018). Complementary to these papers, we examine the joint distribution of markups, markdowns, and other firm characteristics across a wide cross-section of the economy. A notable finding from this analysis is evidence for Marshall's *anti*-Second Law: the elasticity of demand increases with quantity.

Our finding in support of Marshall's *anti*-Second Law, which states that price elasticity increases with quantity demanded, contributes to two related strands of literature in trade and macroeconomics. The first strand examines how market size affects entry, competition, and welfare, including studies such as Mankiw and Whinston (1986), Melitz (2003), Edmond, Midrigan and Xu (2015), Melitz (2018), Mrázová and Neary (2017, 2020), Matsuyama and Ushchev (2023*b*), and Baqaee, Farhi and Sangani (2024*b*), among others. A key assumption underlying many predictions in these models is Marshall's Second Law of Demand. For instance, Matsuyama and Ushchev (2023*b*) shows that in the generalized class of models of Melitz (2003), Marshall's Second Law of Demand implies that lower entry costs intensify competitive pressures, reducing markups for all firms. Similarly, larger market sizes lead to lower markups across firms. Mrázová and Neary (2020) highlights that under the second law, an increase in scale amplifies profits for large firms, reinforcing the "Matthew Effect." Our empirical evidence suggests that the demand system may, in fact, exhibit Marshall's *anti*-Second Law of Demand while the supply system may instead exhibit a form of Marshall's Second Law of *Labor Supply*. The macroeconomic implications of this pattern remain unexplored.

Empirical support for Marshall's Second Law of Demand at a macroeconomic scale primarily arises from observed incomplete price pass-through: firms do not fully transmit cost or total factor productivity (TFP) shocks into prices. Under the common assumption of competitive input markets, incomplete price pass-through implies that markups increase with quantity demanded, meaning that larger firms tend to have higher markups in many macroeconomic models. This interpretation underlies findings in studies such as Feenstra, Gagnon and Knetter (1996), Nakamura and Zerom (2010), Goldberg and Hellerstein (2013), and more recently, Amiti, Itskhoki and Konings (2019) and Sangani (2023). Many empirical studies infer markup behavior by examining how input price shocks pass through to final goods prices. However, we show theoretically and quantitatively that observed incomplete pass-through may instead result from an upward-sloping supply curve or the presence of monopsony power in labor markets. Distinguishing between these mechanisms is crucial, as different sources of incomplete pass-through have significant implications for macroeconomic dynamics, including price and wage rigidity, unemployment fluctuations, and dividend volatility.

Our findings contribute to the literature on the interaction between labor market power and aggregate economic outcomes. Recent studies have explored how labor market power affects wage setting, the labor share, employment dynamics, and economic inequality (e.g. Webber, 2015; Lipsius, 2018; Berger, Herkenhoff and Mongey, 2022; Lamadon, Mogstad and Setzler, 2022; Benmelech, Bergman and Kim, 2022; Hurst et al., 2022; Berger et al., 2023; Seegmiller, 2023). While much of this literature has focused on monopsony power in labor markets in isolation, we provide evidence on how labor market power varies systematically across firms and interacts with product market power. Our findings suggest that understanding the cross-sectional distribution of markups, markdowns, and firm characteristics is important for interpreting long-run trends in the labor income distribution and firm dynamics.

Finally, our modeling approach aligns with an emerging literature that generalizes and moves beyond the constant elasticity of substitution (CES) demand system (e.g Matsuyama and Ushchev, 2017, 2023*b*; Grossman, Helpman and Lhuillier, 2023; Baqaee, Farhi and Sangani, 2024*a*,*b*). Specifically, we adopt the class of homothetic single-aggregator (HSA) systems first introduced by Matsuyama and Ushchev (2017), and subsequently used by the aforementioned papers. However, our contribution extends beyond previous applications by employing this framework not only for the demand system but also for the labor supply system, allowing us to analyze a monopolistic-monopsonistic structure. These homothetic systems nest CES, translog, and a variety of other demand specifications. By specifying these systems non-parametrically, we derive general theoretical results while also flexibly accommodating data-implied patterns in our calibration.

Outline. Section 2 describes the data. Section 3 outlines the price markup and wage markdown estimation procedure. Section 4 documents and examines the aggregate trends and firm-level cross-sectional characteristics. Section 5 develops a model to rationalize the documented empirical findings and Section 6 analyzes various quantitative exercises with the model. Finally, Section 7 concludes.

2 Data

Our data analysis relies primarily on two main sources. The first is the Longitudinal Business Database (LBD), which a restricted-use business establishment dataset from the U.S. Census Bureau (United States Census Bureau, 2022).[5] The second main dataset is the CRSP/Compustat Merged database (Center for Research in Security Prices, 2020). This dataset is provided by the Wharton Research Data Services (WRDS). We merge these datasets together to form a firm-year level panel that includes both financial variables as well as payroll and employment data. We also utilize

[5]We would like to thank Bryan Seegmiller, Lawrence Schmidt, Dimitris Papanikolaou, and Jonathan Rothbaum for providing access to these data.

various standard macroeconomic and financial time series as auxiliary data. The remainder of this section provides more detail on the LBD (Section 2.1) and discusses the merged dataset (Section 2.2). Appendix B provides more information on the data and variable construction.

2.1 Longitudinal Business Database

The LBD is a census of the U.S. private sector (non-farm) at the establishment level. The database is created using information from the Business Register (BR), formally known as the the Standard Statistical Establishment List (SSEL). This registry is continuously updated through surveys and other administrative records such as those from the IRS. The LBD provides an annual snapshot of various firm dynamics such as the opening or closure of establishments as well as their sales and employment. Establishments in the LBD have both establishment identifiers and firm identifiers, which allow us to aggregate to the firm-level. Our version of the data has coverage from 1976 to 2019. Chow et al. (2021) provide more information on the construction of the LBD.

The most important variables that we utilize from this database are the payroll and employment variables. While Compustat does provide firm-level employment (EMP) and total employee compensation (XLR), its coverage, especially for payroll, is highly incomplete. The LBD, therefore, is able to provide more comprehensive and consistent information on a publicly traded firm's U.S. employment and payroll.

2.2 Merged CRSP/Compustat-LBD Sample

We link the aggregated LBD dataset to the CRSP/Compustat sample using a crosswalk developed and provided by Lawrence Schmidt.[6] Once the data are merged, we construct various key variables and then we apply the standard filters onto the Compustat variables to remove missing or extreme values as well as and certain industries such as regulated utilities (NAICS code 22), financial services (NAICS code 52), real estate (NAICS code 53), public administration (NAISCS code 92), and nonclassifable (NAICS code 99). Appendix B provides more information on this procedure. The final dataset used for the analysis ranges from 1977 to 2019.

Tables 1 and 2 show the summary statistics of firm characteristics of the final merged sampled used in the analysis and the original Compustat sample over the same year range, respectively. Note that in Table 2 we do not include statistics on the wage bill, employment, and labor share since these variables cannot be constructed for all observations for this sample. The firms in the final sample are somewhat larger than the firms in the original Compustat sample as measured by standard size metrics such as sales, total assets, market capitalization, or capital stocks.

[6]We thank Lawrence Schmidt and Brice Green for sharing the crosswalk.

Table 1: Final Sample Summary Statistics

Variable	Mean (1)	SD (2)	P10 (3)	P25 (4)	Median (5)	P75 (6)	P90 (7)	Obs. (8)
Log Sales	19.820	2.042	17.210	18.350	19.780	21.220	22.550	69,500
Log COGS	19.330	2.104	16.630	17.820	19.310	20.770	22.100	69,500
Log SGA	18.210	1.936	15.780	16.800	18.120	19.500	20.810	69,500
Log Wage Bill	18.070	1.938	15.590	16.700	18.080	19.390	20.610	69,500
Log Employment	7.179	2.006	4.605	5.793	7.211	8.561	9.801	69,500
Log Physical Capital	18.160	2.358	15.160	16.470	18.070	19.770	21.340	69,500
Log Intangible Capital	18.260	2.000	15.770	16.810	18.140	19.540	20.960	69,500
Log Total Assets	19.670	2.097	16.990	18.130	19.570	21.100	22.470	69,500
Log Market Cap	19.020	2.397	16.000	17.230	18.920	20.700	22.170	69,000
Labor Share VA	0.668	0.391	0.298	0.470	0.641	0.789	0.956	69,500

Notes: This table presents the summary statistics of the final sample used in the analysis. The sample ranges from 1977 to 2019. All nominal variables are deflated using the BEA's GDP Price Deflator. Column (1) reports the mean, and Column (2) reports the standard deviation. Columns (3) to (7) report the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile, respectively. Column (8) reports the number of observations. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table 2: Full Compustat Sample Summary Statistics

Variable	Mean (1)	SD (2)	P10 (3)	P25 (4)	Median (5)	P75 (6)	P90 (7)	Obs. (8)
Log Sales	19.040	2.484	15.900	17.410	19.110	20.720	22.200	151,000
Log COGS	18.490	2.577	15.210	16.790	18.550	20.250	21.770	151,000
Log SGA	17.650	2.098	15.050	16.140	17.550	19.020	20.440	151,000
Log Physical Capital	17.580	2.692	14.190	15.660	17.510	19.440	21.190	151,000
Log Intangible Capital	17.750	2.150	15.080	16.200	17.640	19.130	20.620	151,000
Log Total Assets	19.080	2.396	16.040	17.390	19.050	20.710	22.230	151,000
Log Market Cap	18.640	2.522	15.510	16.810	18.530	20.380	21.980	140,000

Notes: This table presents the summary statistics of the original Compustat sample. The sample ranges from 1977 to 2019. All nominal variables are deflated using the BEA's GDP Price Deflator. Column (1) reports the mean, and Column (2) reports the standard deviation. Columns (3) to (7) report the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile, respectively. Column (8) reports the number of observations. All figures are rounded in accordance with U.S. Census disclosure requirements.

3 Estimating Price Markups and Wage Markdowns

This section describes the framework to estimate firm-level price markups and wage markdowns. The approach follows most closely that of Hall (1988), De Loecker and Warzynski (2012), Dobbelaere and Mairesse (2013) and Yeh, Macaluso and Hershbein (2022). We define price markups and wage markdowns for firm i in period t as follows

$$\mu_{i,t} = \frac{P_{i,t}}{\text{MC}_{i,t}}, \tag{1}$$

$$\nu_{i,t} = \frac{\text{MRPL}_{i,t}}{W_{i,t}}, \tag{2}$$

where the price markup $\mu_{i,t}$ is the ratio between the price of output $P_{i,t}$ and its marginal cost $\text{MC}_{i,t}$ and the wage markdown $\nu_{i,t}$ is the ratio between the marginal revenue product of labor $\text{MRPL}_{i,t}$ to the wage paid $W_{i,t}$.[7] A value of unity for either measure implies no market power in that market whereas a value greater than one implies that firm has market power in the respective market.

Both the firm's price markup and wage markdown can be characterized by the firm's optimality conditions from either the profit maximization or cost minimization problem. Thus, one can relate markups and markdowns to the ratio of output elasticities to cost shares. Therefore, we can utilize what is commonly referred to as ratio estimators to recover firm markups and markdowns. Output elasticities and cost shares are either directly observable or easier to estimate than directly recovering marginals costs or marginal revenue products. With this insight, we estimate the firm's production function following the method of De Loecker and Warzynski (2012). The remainder of this section discusses the ratio estimators (Section 3.1) and the production function estimation (Section 3.2).

3.1 Deriving Markups and Markdowns from the Firm's Problem

We start with a general but simple static firm profit maximization problem to derive the relationship linking markups and markdowns with output elasticities and cost shares. Consider the profit maximization problem of a firm that uses $J > 1$ inputs to produce one output, which is given by

$$\begin{gathered} \max_{\mathbf{X}_{i,t} \in \mathbb{R}^J_{++}} P_{i,t}(Y_{i,t})Y_{i,t} - \sum_{j=1}^{J} W^j_{i,t}(X^j_{i,t})X^j_{i,t} \\ \text{subject to} \\ Y_{i,t} \leq F(\mathbf{X}_{i,t}; \omega_{i,t}), \end{gathered} \tag{3}$$

where $\mathbf{X}_{i,t}$ is the vector of inputs, $P_{i,t}(\cdot)$ is the inverse demand function, $Y_{i,t}$ is output, $W^j_{i,t}(X^j_{i,t})$ is the inverse supply function of input j, $F(\cdot;\cdot)$ is the production function that satisfies the usual assumptions, and $\omega_{i,t}$ is the firm's Hicks-neutral log productivity. Finally, let $\mathcal{J}$ denote the set of inputs. Note that we assume that the inverse demand function of a given input j only depends on

[7] While the definition of the price markup is standard, some papers define the wage markdown as the ratio of wages to MRPL instead (e.g. Berger, Herkenhoff and Mongey, 2022). We follow the convention of defining the wage markdown as the ratio of MRPL to wages for ease of exposition in our context.

the quantity of that input.[8]

In order to solve the above model and recover expressions for the price markup and wage markdown (or more generally markdowns for any input in which the firm has market power over), we at need at least one input that is flexible. Let $\mathcal{F} \subseteq \mathcal{J}$ denote the set of flexible inputs. A flexible input satisfies all of the following assumptions:

Assumption 1. *Inputs $f \in \mathcal{F}$ have no adjustment costs.*

Assumption 2. *Inputs $f \in \mathcal{F}$ are not subject to monopsony or oligopsony forces, i.e. $W^f_{i,t}(X^f_{i,t}) = \bar{W}^f_{i,t} \in \mathbb{R}_{++}$, that is the firm takes the input price as given.*

Assumption 3. *Inputs $f \in \mathcal{F}$ are chosen statically.*

Assumption 4. *The production function $F(\cdot;\omega_{i,t})$ is twice continuously differentiable with respect to $X^f_{i,t}$ and satisfies*

$$\lim_{X^f_{i,t}\to 0} \frac{\partial F(\mathbf{X}_{i,t};\omega_{i,t})}{\partial X^f_{i,t}} = +\infty,$$
$$\lim_{X^f_{i,t}\to +\infty} \frac{\partial F(\mathbf{X}_{i,t};\omega_{i,t})}{\partial X^f_{i,t}} = 0,$$

for all $f \in \mathcal{F}$ and for all $\omega_{i,t} \in \mathbb{R}$. Finally, the firm's inverse demand function for f, $P_{i,t}(\cdot)$, is continuously differentiable and strictly decreasing.

Assumption 5. *Inputs $f \in \mathcal{F}$ are only used for the production of output only.*

With these assumptions, we can derive the estimator for the firm's price markup which then leads us to the estimator for the firm's wage markdown. Proposition 1 introduces the ratio estimator for price markups.

Proposition 1 (Price Markup Ratio Estimator)**.** *If $\mathcal{F}$ is non-empty, that is there exists at least one $f \in \mathcal{F}$, then the firm i's price markup can be characterized by*

$$\mu_{i,t} = \frac{\theta^f_{i,t}}{\alpha^f_{i,t}}, \tag{4}$$

where $\theta^f_{i,t}$ is the firm's output elasticity with respect to input f and $\alpha^f_{i,t}$ is the share of revenue of input f.

Proof. See Appendix C.1 for the proof. □

[8] The model can account for oligopolistic/oligopsonistic competition in a Cournot setting in which the firm takes its competitors' choices as given without materially changing the model. We also assume there is no collusion in any market for simplicity. However, the method we use to estimate production functions and markups/markdowns does not impose any assumptions on conduct. See Delabastita and Rubens (2024) for a further discussion and a study that examines the contribution of the elasticity of supply and collusion on wage markdowns.

Notice that in the proof of Proposition 1 we also derived the relationship between price markups and the elasticity of demand. With Proposition 1, we can also now define the ratio estimator for wage markdowns.

Proposition 2 (Input Markdown Ratio Estimator)**.** *Suppose there exists at least one $f \in \mathcal{F}$. Therefore, the results from Proposition 1 hold for inputs $f \in \mathcal{F}$. Let $\mathcal{L}$ be the set of inputs that satisfy Assumptions 1 and 3 to 5. Then firm i's markdown for input $j \in \mathcal{L}$ is given by*

$$\nu^j_{i,t} = \frac{\theta^j_{i,t}}{\alpha^j_{i,t}} \mu^{-1}_{i,t}. \tag{5}$$

Proof. See Appendix C.1 for the proof. □

Since Propositions 1 and 2 rely on output elasticities, we need to estimate production functions in order to estimate markups and markdowns using the ratio estimators. Section 3.2 discusses this procedure in greater detail. There are further adjustments that can be made to the ratio estimators if some of the above assumptions are violated (such as accounting for adjustment costs or non-production inputs). These are discussed in Appendices C.2 and C.3.[9]

3.2 Production Function Estimation

We follow the production function estimation procedure developed by De Loecker and Warzynski (2012), which relies on a proxy variable approach and builds upon the methods of Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves and Frazer (2015). The estimation consists of two stages, the first stage recovers expected output and an error term and the second stage estimates the production function parameters given the productivity process. This section gives a brief outline of this approach. We also discuss various steps we take to address concerns with this estimation approach. Appendix D goes into greater detail on how we implement the production function estimation, particularly on the second-stage.

Let lowercase letters denote logs; thus, we can express the firm's log output as

$$y_{i,t} = f(\mathbf{x}_{i,t}; \boldsymbol{\beta}) + \omega_{i,t} + \varepsilon_{i,t}, \tag{6}$$

where $f(\cdot;\cdot)$ is the log production function, $\mathbf{x}_{i,t}$ is the vector of log inputs, $\boldsymbol{\beta}$ is the vector of parameters, and $\varepsilon_{i,t}$ is the error term. We estimate a translog production function, which can approximate any arbitrary differentiable production function to a second-order. A translog function production is given by

$$f(\mathbf{x}_{i,t}; \boldsymbol{\beta}) = \sum_{j \in \mathcal{J}} \beta_j x^j_{i,t} + \frac{1}{2} \sum_{j \in \mathcal{J}} \sum_{j' \in \mathcal{J}} \beta_{j,j'} x^j_{i,t} x^{j'}_{i,t}. \tag{7}$$

[9] Yeh, Macaluso and Hershbein (2022) and Bond et al. (2021) also discuss these adjustments.

We estimate a production function that is time invariant for each two-digit NAICS industry (NAICS2) in the sample. Furthermore, in our implementation we use four inputs (in logs): labor $l_{i,t}$, materials $m_{i,t}$, physical capital $k_{i,t}$, and intangible capital $n_{i,t}$.[10]

Labor, physical capital, and intangible capital are constructed in a straightforward manner. Labor is taken directly from the LBD. Physical capital is computed using a standard capitalization approach and we also compute intangible capital using the capitalization approach from Eisfeldt and Papanikolaou (2013) and Peters and Taylor (2017).[11]

Materials is constructed as follows

$$M_{i,t} = \text{COGS}_{i,t} + \text{XSGA}_{i,t} - \text{WB}_{i,t} - X_{i,t}. \tag{8}$$

We remove the wage bill and other fixed costs $X_{i,t}$ (such as rents and the capitalized portion of intangibles) from the sum of $\text{COGS}_{i,t}$ (cost of goods sold) and $\text{XSGA}_{i,t}$ (selling and general administration expenses; SGA). We define materials following Equation (8) to address concerns with using only COGS to estimate markups.[12] De Loecker, Eeckhout and Unger (2020) also propose a similar approach; however, they do not implement this for the Compustat sample as it is not feasible to implement using Compustat alone, as labor expenses are not reliably reported. This approach assumes that labor expenses (and other expenses such as rent) represent a sufficiently large component of the non-flexible inputs and that the residual is a much better approximation of a flexible input that can be used to estimate markups.

Once we have created the above variables, we can run the first stage of the estimation procedure. We represent the firm's output as

$$y_{i,t} = \phi_t(\mathbf{x}_{i,t}, \mathbf{z}_{i,t}) + \varepsilon_{i,t},$$

where $\mathbf{z}_{i,t}$ is a vector of controls such as time fixed effects and

$$\phi_t(\mathbf{x}_{i,t}, \mathbf{z}_{i,t}) = \mathcal{P}_k(\mathbf{x}_{i,t}) + h_t(m_{i,t}, k_{i,t}, \mathbf{z}_{i,t}),$$

where $\mathcal{P}_k(\cdot)$ is k-order polynomial function and $h_t(\cdot,\cdot,\cdot)$ is the control function. We implement the first stage by regressing $y_{i,t}$ onto a second-order polynomial of $\mathbf{x}_{i,t}$ with year fixed effects via OLS. We can then recover the predicted output $\hat{\phi}_{i,t}$ and predicted error terms $\hat{\varepsilon}_{i,t}$, which are utilized in the second stage.

With the predicted output values, we can recover an estimate of firm-level productivity as follows

$$\omega_{i,t}(\hat{\boldsymbol{\beta}}) = \hat{\phi}_{i,t} - f(\mathbf{x}_{i,t}; \hat{\boldsymbol{\beta}}),$$

[10] We assume that $\beta_{j,j'} = \beta_{j',j}$ for all $j, j' \in \mathcal{J}$. Therefore, if there are J inputs, then with a translog production function $\boldsymbol{\beta}$ is of dimension $2J + \binom{J}{2} = J(J+3)/2$, which consists of J first-order terms, J second-order terms, and $\binom{J}{2}$ cross terms.

[11] Appendix B.2 contains more details on how these inputs are constructed.

[12] See Traina (2018) for a more thorough discussion of the issues of using COGS as a flexible input.

where $\hat{\boldsymbol{\beta}}$ is the estimate of the production function coefficients. Firm-level productivity follows the law of motion

$$\omega_{i,t}(\hat{\boldsymbol{\beta}}) = g(\omega_{i,t-1}(\hat{\boldsymbol{\beta}})) + \xi_{i,t}(\hat{\boldsymbol{\beta}}),$$

where $\xi_{i,t}(\hat{\boldsymbol{\beta}})$ is the uncorrelated idiosyncratic productivity shock and $g(x) = (1 - \rho_\omega)\bar{\omega} + \rho_\omega x$, where $\bar{\omega}$ is the unconditional mean and $\rho_\omega \in (-1, 1)$. Given $\hat{\boldsymbol{\beta}}$, $\xi_{i,t}(\hat{\boldsymbol{\beta}})$ is recovered by regressing $\omega_{i,t}(\hat{\boldsymbol{\beta}})$ onto its lag and a constant.

The estimates $\hat{\boldsymbol{\beta}}$ are obtained via GMM with the following moment condition

$$\mathbb{E}\left[\xi_{i,t}(\hat{\boldsymbol{\beta}}) \cdot \tilde{\mathbf{z}}_{i,t}\right] = \mathbf{0}, \tag{9}$$

where $\tilde{\mathbf{z}}_{i,t}$ is a vector of instruments that has a dimension of at least that of $\boldsymbol{\beta}$. These instruments are the current physical and intangible capital, lagged values of labor and material, and interactions of these. The identification assumption here is that the firm makes its materials, labor, and capital investment decisions after observing $\xi_{i,t}(\boldsymbol{\beta})$. With the estimates of $\boldsymbol{\beta}$, output elasticities can be computed in a straightforward manner. Since we have a translog production function all output elasticities are linear functions of $\mathbf{x}_{i,t}$ with coefficients $\hat{\boldsymbol{\beta}}$. From this we can compute the markups and markdowns following Propositions 1 and 2.[13] We also can recover an estimate of the firm's log productivity $\omega_{i,t}$. However, since we are estimating revenue-deflated production functions, the productivity term recovered is the log of revenue-based total factor productivity (TFPR).

Before examining the results, we first address and discuss various concerns and issues with our approach.[14] We also discuss mitigation strategies and caveats.

Price Markup Level. The first issue is that since we do not observe output prices and the prices of some inputs, we are relying on a revenue-deflated approach. The issue is that using revenue-deflated inputs introduces a bias to the estimated output elasticity and thus bias the ratio estimators. Under certain circumstances the markup estimate may even be completely uninformative (Bond et al., 2021). However, De Ridder, Grassi and Morzenti (2025) show theoretically and empirically that under circumstances in which our empirical setting satisfies, the markup estimate is informative in terms of its time series and cross-sectional information–while there is a multiplicative bias of unknown direction in the estimated level of markups, the relative ordering is preserved. Thus, covariances with respect to other variables and the direction of time series trends are preserved for markups.

Wage Markdown Level. The second concern regards the level of markdown: given the unidentified level of markup, one might worry that the level of markdown is also unidentified. However, Yeh, Macaluso and Hershbein (2022) show that the markdown ratio estimator is immune to this

[13] We also apply the correction term following De Loecker and Warzynski (2012) to compute expenditure shares. This approaches removes any output variation not related to variables impacting input demand and market characteristics.

[14] Appendix C contains more detail on these issues as well as others not discussed here.

issue, as the bias enters as a constant multiplicative factor across all output elasticities and therefore cancels out, yielding a consistent estimate of markdown levels. Yet another important concern is the existence of demand-shifting labor—workers who do not contribute directly to production but engage in activities that affect demand, such as marketing—which is a non-negligible issue in our Compustat sample. In Appendix C.3, we show that demand-shifting labor can bias the level of markdowns. Intuitively, an unknown portion of demand-shifting labor deflates the labor output elasticity since these workers do not enter the production process; however, this effect creates a bias in markdown levels. Because of these biases, the existing literature also finds that markdown levels vary widely, with markdown sometimes falling below 1.[15] Based on these concerns, we focus most of our analysis on the relative ordering and ranking of markups and markdowns rather than their absolute levels.

Identification. The next concern is that, while these moment conditions are relatively straightforward and sensible, Gandhi, Navarro and Rivers (2020) show that point identification relies on time-series variation in flexible input prices and may be problematic in short time-series settings. Although this issue can arise in settings with shorter time window, our empirical context benefits from a long panel, which helps mitigate it. Nonetheless, we implement the solution suggested by Flynn, Traina and Gandhi (2019) and impose a constant returns to scale (CRS) restriction on our moment conditions in (9).[16] This assumption is reasonable given substantial evidence for constant or slightly decreasing returns to scale (e.g., Basu and Fernald, 1997; Syverson, 2004*a*,*b*). We estimate the production function both with and without the CRS restriction as a robustness check, and we present the latter results in Appendix A. Our main specification uses the estimates that impose the CRS restriction.

Existence of Monopsony Power over Other Inputs. The fourth major concern is the construction of the flexible input, which we have already briefly discussed. One remaining issue is that if the input chosen to be used as the flexible input for the markup estimate violates Assumption 2 (no monopsony or oligopsony power). If this assumption is violated then using the estimator in (4) results in

$$\frac{\theta_{i,t}^j}{\alpha_{i,t}^j} = \nu_{i,t}^j \mu_{i,t}, \tag{10}$$

which follows directly from Proposition 2. Equation (10) shows that using an input j that violates Assumption 2 to estimate price markups can lead to confounding trends in the markdown of input j to trends in the price markup. This result holds both for individual firm-level results and for the aggregate time series. Furthermore, since the markup is used to estimate markdowns, we would be underestimating the markdown for the other inputs. However, we believe that our method of

[15] See Traina (2018), Yeh, Macaluso and Hershbein (2022), Mertens and Mottironi (2023) for examples of this.

[16] Appendix D discusses implementation details.

purging the sum of COGS and SGA of labor expenses and other fixed costs is able to remove most of this bias. We examine this issue more closely when discussing the results in Section 4.

Factor Augmenting Productivity. The last major concern is that our production estimation procedure does not allow for factor augmenting productivity or factor-biased technological change. Doraszelski and Jaumandreu (2019), Demirer (2022), and Raval (2023*a*,*b*) show that not accounting for labor augmenting productivity will bias the estimates of output elasticity. However, the proposed solutions require data (such as quantity data) which is not available in our setting and/or do not allow the researcher to also recover labor market power.

4 Empirical Results

4.1 Aggregate Trends

We calculate aggregate price markups and wage markdowns as follows

$$\bar{\mu}_t = \sum_{i \in \mathcal{S}_t} s_{\mu,i,t} \mu_{i,t}, \tag{11}$$

$$\bar{\nu}_t = \sum_{i \in \mathcal{S}_t} s_{\nu,i,t} \nu_{i,t}, \tag{12}$$

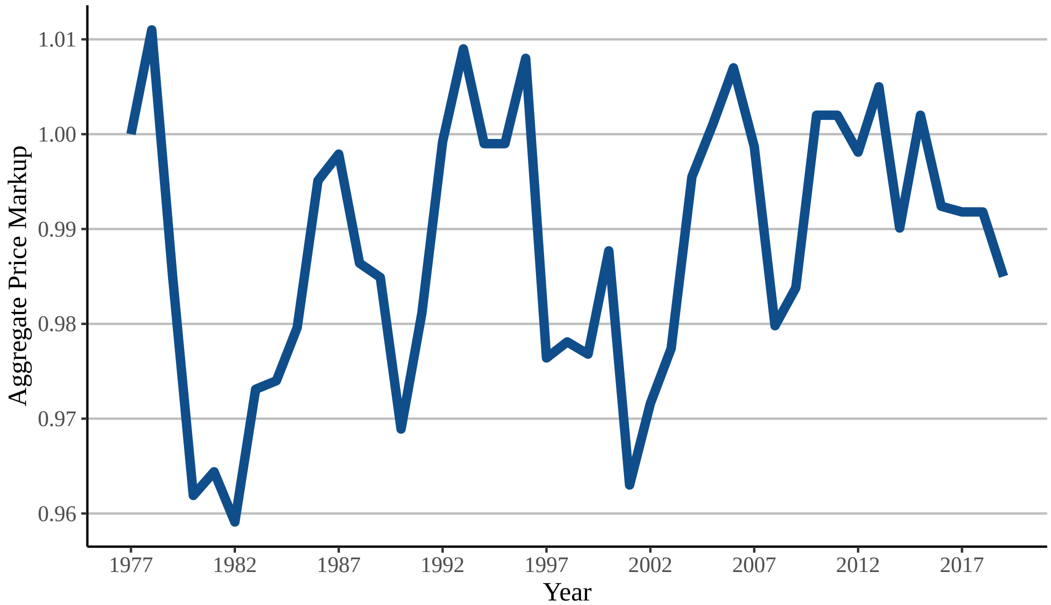
where $\mathcal{S}_t$ is the set of firms observed in year t, and $s_{\mu,i,t}$ and $s_{\nu,i,t}$ denote the weights for the aggregate markup and markdown, respectively. In our baseline specification, we use the CRS-restricted estimates for both markups and markdowns, weighting by sales for the aggregate markup and by the wage bill for the aggregate markdown. Figures 1 and 2 plot the evolution of these indices from 1977 to 2019.[17] Because the absolute levels of markup and markdown may be unidentified as we discussed in the last section, we report index versions normalized to one in 1977; this preserves all trends and cross-sectional rankings and covariances. Details on our index construction, which is slightly more involved than simply normalizing the 1977 value, are provided in Appendix B.

The aggregate markup remains relatively stable throughout the period, in contrast to the sharply rising trend reported by De Loecker, Eeckhout and Unger (2020). By contrast, the aggregate markdown exhibits a significant upward trend. It rises steadily from 1977 until the early 1990s, followed by a sharp increase of over 45% between 1993 and 1997. After a moderate decline of roughly 11% in the mid-2000s, the markdown resumes its upward trajectory, increasing by over 30% by 2019, with a modest dip in the mid-2010s.

Despite differing from De Loecker, Eeckhout and Unger (2020), our findings align with prior other work. For example, Anderson, Rebelo and Wong (2018) show that the aggregate markup in the retail sector in both the United States and Canada has remained relatively stable over this period. To understand why our result for price markups diverges from De Loecker, Eeckhout

[17] See Figures A1 and A2 in Appendix A.1 for the corresponding level-based series.

Figure 1: Evolution of the Aggregate Price Markup (Index), 1977–2019



Notes: This figure reports the aggregate price markup following the definition in Equation (11) from 1977 to 2019 as an index. Production functions are estimated using a CRS restriction, and the aggregate price markup is computed using sales as the weight. All figures are rounded in accordance with U.S. Census disclosure requirements.

and Unger (2020), consider a simple example in which the measured flexible input is entirely labor subject to monopsony power. From Proposition 2, one can show that the measured markup combines both markup and markdown:

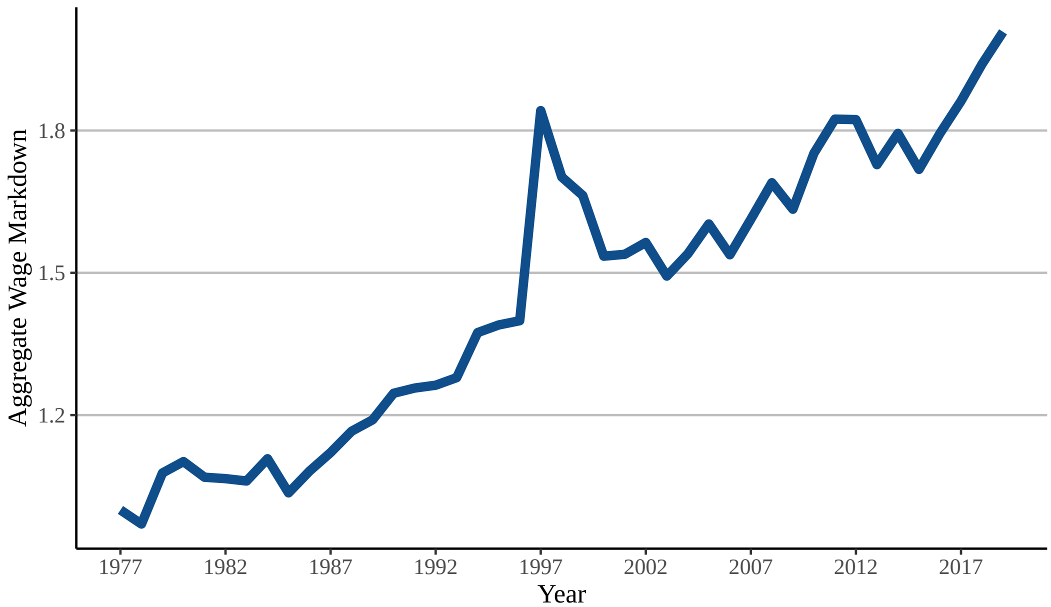
$$\log \mu_{i,t}^{\text{measured}} = \log \mu_{i,t} + \log \nu_{i,t}, \tag{13}$$

Given how we construct inputs and estimate market power, our results can be reconciled with those of De Loecker, Eeckhout and Unger (2020). Whereas De Loecker, Eeckhout and Unger (2020) use COGS as the flexible input, which includes labor expenses, our approach separates product and labor inputs; see Section 4.2 for details. Therefore, earlier findings of rising measured markups may actually reflect an increase in total market power (the overall wedge) rather than a rise in price markups alone.

Comparing our markdown results to the literature, both Yeh, Macaluso and Hershbein (2022) and Traina (2018) estimate wage markdowns for the U.S. manufacturing sector, finding increases of roughly 10% and 100%, respectively. We show that, across a broader set of industries and publicly traded U.S. firms, markdowns have increased substantially and are generally closer to the magnitude reported by Traina (2018) in manufacturing.[18]

[18]We observe similar industry-level patterns; see Table A1 in Appendix A.2 for NAICS2-level results.

Figure 2: Evolution of the Aggregate Wage Markdown (Index), 1977–2019



Notes: This figure reports the aggregate wage markdown following the definition in Equation (12) from 1977 to 2019 as an index. Production functions are estimated using a CRS restriction, and the aggregate wage markdown is computed using the wage bill as the weight. All figures are rounded in accordance with U.S. Census disclosure requirements.

We verify robustness by re-estimating both the aggregate markup and markdown under alternative weighting schemes and with versus without the CRS restriction, following Edmond, Midrigan and Xu (2023). Figures A3 and A4 compare CRS and no-CRS markups under sales-weighted, harmonic sales-weighted, and unweighted averages. In every case the aggregate markup remains essentially flat. Likewise, Figures A5 and A6 show CRS and no-CRS markdowns under wage-bill, employment, and unweighted averages: all specifications exhibit the same upward trend, with wage-bill weights producing the highest levels (followed by employment and then unweighted), consistent with larger firms paying higher wages (e.g., Brown and Medoff, 1989; Bloom et al., 2018). Finally, adjusting the markdown for demand-shifting labor (e.g., advertising or R&D), as detailed in Appendix C.3, raises its level slightly but leaves the overall trend intact (Figure A11).

4.2 Decomposing the Aggregate Trends

We decompose the time series trends in the aggregate price markup and wage markdown across two dimensions to understand their underlying drivers. First, given the ratio estimators used to estimate markups and markdowns, we can decompose these series to recover the aggregate output elasticity and cost share of materials and labor and analyze how these components evolve over time. Second, we examine how the cross-sectional dispersion changes over time and how that affects the aggregate trends.

Recall the ratio estimators from Equations (4) and (5). We can construct aggregate analogs of the output elasticity and cost share given our aggregate markup and markdown estimates.[19] Notice that there is a sensible aggregate analog of cost shares (simply the sum of the total expenditure of that input divided by total sales) and that it is directly observable. Thus, along with our aggregate markup and markdown estimates, we can back out a corresponding aggregate elasticity. Let $\bar{\theta}_t^f(\bar{\mu}_t, \bar{\alpha}_t^f)$ be the function that admits the implied aggregate output elasticity of materials given an aggregate markup estimate and the cost share of materials. This is given by

$$\bar{\theta}_t^f(\bar{\mu}_t, \bar{\alpha}_t^f) = \bar{\mu}_t \bar{\alpha}_t^f, \tag{14}$$

which follows from rearranging Equation (4) to isolate for the aggregate output elasticity of the flexible input. The aggregate output elasticity of materials depends on the chosen aggregate markup specification. Following a similar approach, let $\bar{\theta}_t^l(\bar{\nu}_t, \bar{\alpha}_t^l, \bar{\mu}_t)$ denote the function to produce the implied aggregate output elasticity of labor, which is given by

$$\bar{\theta}_t^l(\bar{\nu}_t, \bar{\alpha}_t^l, \bar{\mu}_t) = \bar{\nu}_t \bar{\alpha}_t^l \bar{\mu}_t, \tag{15}$$

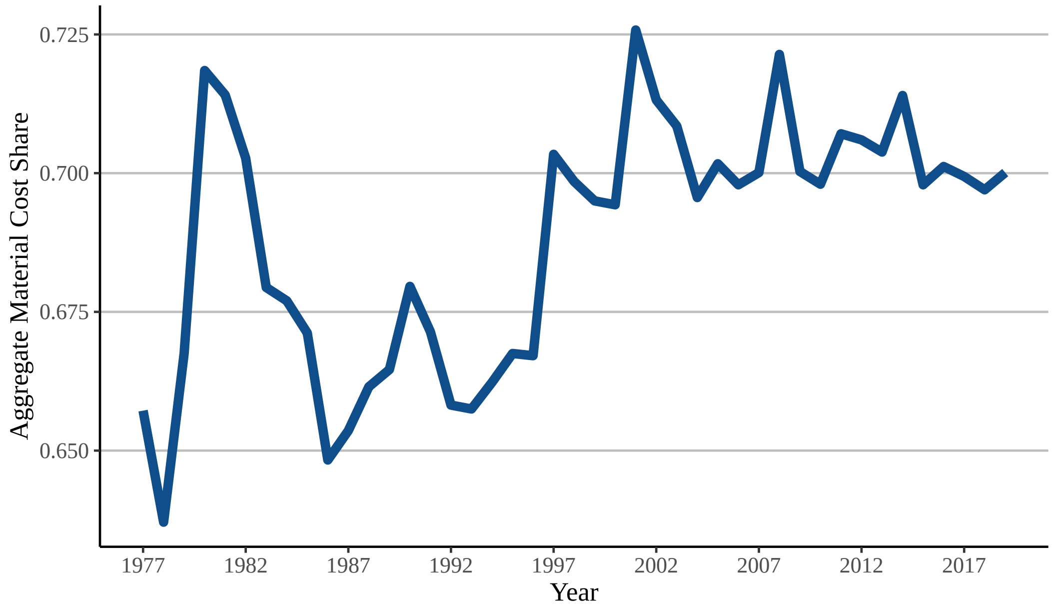
which follows from rearranging Equation (5) to isolate for the output elasticity of labor. Equation (15) shows that the aggregate output elasticity of labor depends not only on the chosen aggregate markdown specification but also the aggregate markup specification.

Figure 3 reports the decomposition following Equation (14) using the baseline specification for the aggregate price markup and the sales-weighted aggregate cost share of materials. Both series move closely together over time. Both series moderately increase in the time period. The aggregate cost share of materials starts at 0.66 in 1977 and increases to just over 0.70 in the late-1990s and stabilizing there. The aggregate output elasticity of materials starts at 0.65 in 1977 and increases to 0.69 in 2019. The series, like the aggregate cost share, stabilizes at around 0.70 in the late-1990s. This stability is consistent with De Loecker, Eeckhout and Unger (2020), who find limited variation in estimated output elasticities over time. Thus, for the aggregate price markup, the time-series pattern can be almost largely matches the movements in the aggregate cost share alone, resulting in a stable aggregate markup.

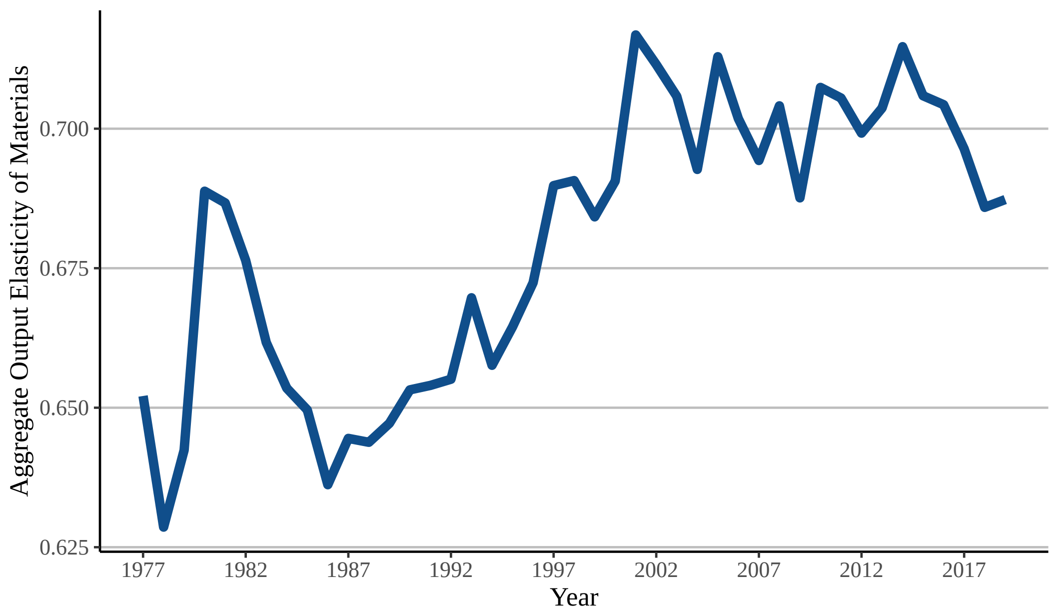
In contrast, the aggregate output elasticity of labor and aggregate cost share of labor evolve across the time period very differently. Figure 4 reports the time series of the aggregate cost share of labor and aggregate output elasticity of labor in Panels (a) and (b), respectively. We report the aggregate output elasticity of labor that is generated by inputting the sales-weighted aggregate cost share of labor, sales-weighted aggregate markup, and wage bill-weighted aggregate markdown

[19] Edmond, Midrigan and Xu (2023) provide a more involved treatment on the aggregation of markups and markdowns. Determining the appropriate weighting methodology to aggregate firm-level estimates in a manner consistent with the concept of an aggregate production function is not always straightforward.

Figure 3: Decomposition of Aggregate Price Markups



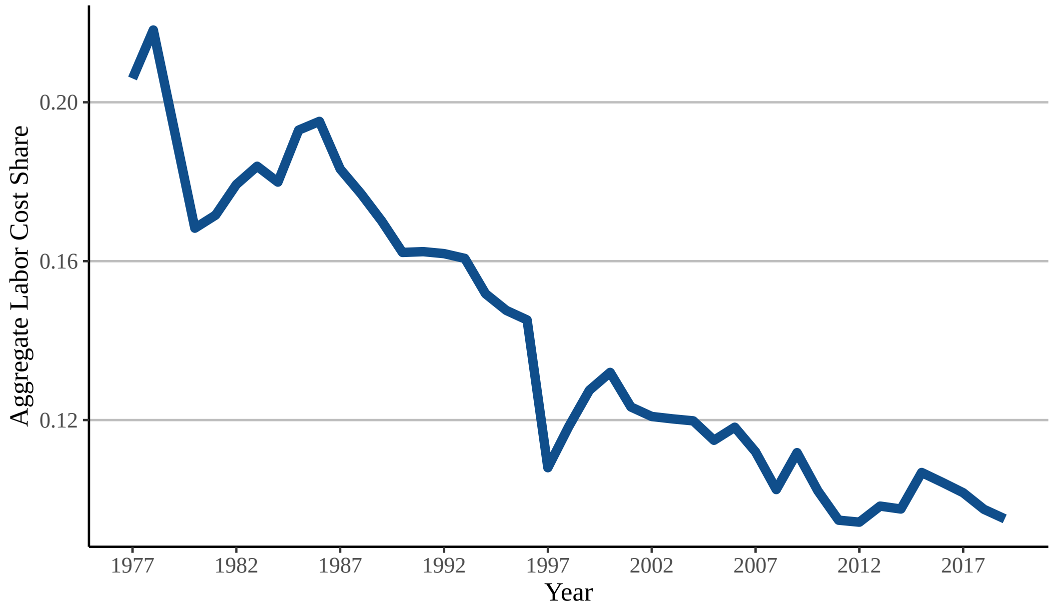
(a) Aggregate Cost Share of Materials



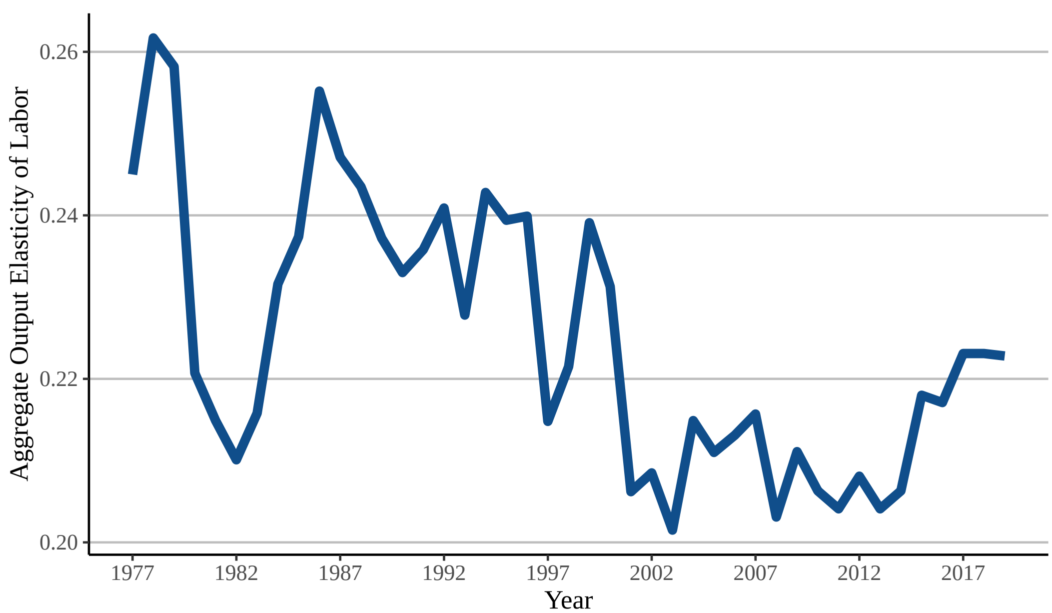
(b) Aggregate Output Elasticity of Materials

Notes: This figure reports the aggregate cost share of materials and the aggregate output elasticity of materials in Panels (a) and (b), respectively. The aggregate cost share is defined as the total materials expenditure divided by total sales in the sample or equivalently the sales-weighted average of the firm-level cost share of materials. The aggregate output elasticity is computed following (14) using the aggregate cost share of materials reported in Panel (a) and the aggregate markup reported in Figure A1. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure 4: Decomposition of Aggregate Wage Markdowns



(a) Aggregate Cost Share of Labor



(b) Aggregate Output Elasticity of Labor

Notes: This figure reports the aggregate cost share of labor and the aggregate output elasticity of labor in Panels (a) and (b), respectively. The aggregate cost share is defined as the total labor expenditure divided by total sales in the sample or equivalently the sales-weighted average of the firm-level cost share of labor. The aggregate output elasticity is computed following (15) using the aggregate cost share of labor reported in Panel (a), the aggregate markdown reported in Figure A2, and the aggregate markup reported in Figure A1. All figures are rounded in accordance with U.S. Census disclosure requirements.

into Equation (15). The aggregate output elasticity of labor declines moderately over time from 0.26 in 1977 to 0.22 in 2019, a 14.6% decline, with a peak decline of 23.0%. The aggregate cost share of labor, however, declines significantly more. It starts in 1977 at 0.21 and declines to 0.10 in 2019, more than halving, which is consistent with Elsby, Hobijn and Şahin (2013), Karabarbounis and Neiman (2014), and Barkai (2020). Hence, the aggregate wage markdown is driven mainly by the falling labor cost share, partly offset by the drop in labor elasticity. Given the magnitude of the labor cost-share decline, which is the dominant driver of our result, our findings are likely robust to alternative (or misspecified) production-function specifications.

These patterns reconcile our results with those of De Loecker, Eeckhout and Unger (2020): they observe a constant COGS elasticity but a declining COGS share, leading to rising markups. Our breakdown in Figures 3a and 4a suggests that the decline in the COGS share largely reflects a decline in the labor share. While the flexible-input elasticity appears flat, this masks offsetting movements in the underlying elasticities. Specifically, the materials elasticity has increased modestly over time, while the labor elasticity has declined.

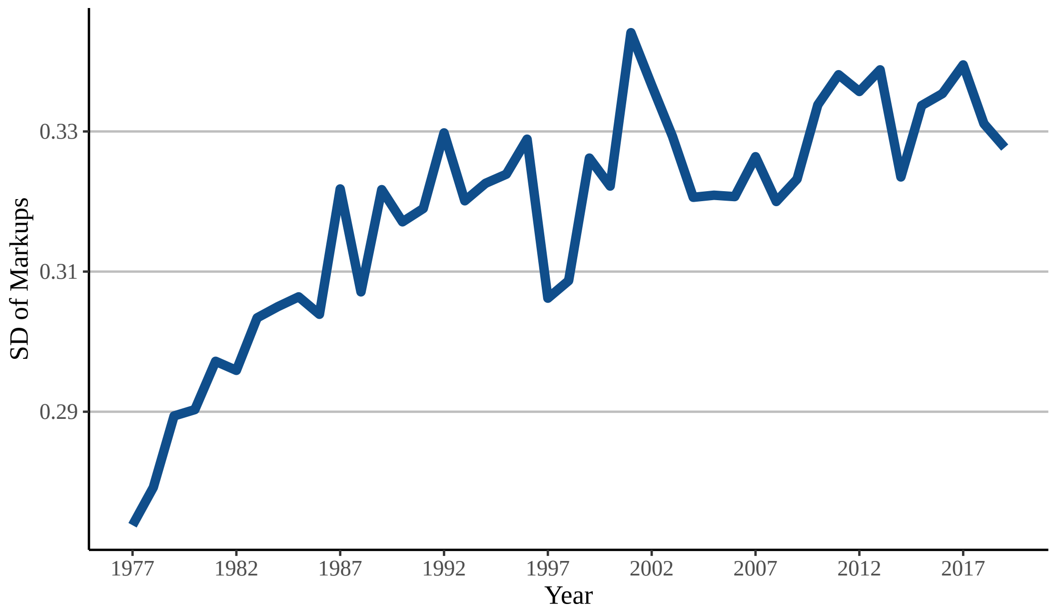
We next turn to the evolution of the cross-sectional distributions of markups and markdowns and analyze their role in driving the aggregate trends. First, we analyze the dispersion in price markups. Figure 5 shows that the standard deviation of markups has grown noticeably (by roughly 20% over the sample) while the gap between the 90th and 50th percentiles has widened by over 50% and the 50th-10th percentile gap by around 23%. The fact that the upper-tail gap expands more rapidly indicates that the highest markup firms are playing an increasingly dominant role in overall dispersion, contributing to greater right-skewness.

The dispersion of wage markdowns evolves similarly over this time period, however, the magnitudes are significantly higher than that of markups. Figure 6 reports the various dispersion measures of wage markdowns. Figure 6a shows that the standard deviation of markdowns increased from 0.84 in 1977 to 2.38 in 2019, which is a growth of over 180%. Both the starting level of the standard deviation of markdowns and the growth rate is significantly higher than those of markups. We also find a similar pattern with the 90/50 and 50/10 differences. Figure 6b reports that the 90/50 difference (blue line) increases from 1.13 in 1977 to 3.37 in 2019 (increase of almost 200%) and that the 50/10 difference (orange line) increases from 0.43 in 1977 to 0.79 in 2019 (increase of 84.3%). Wage markdowns begin with substantially more right-skewness than price markups and exhibits more growth in right-skewness.

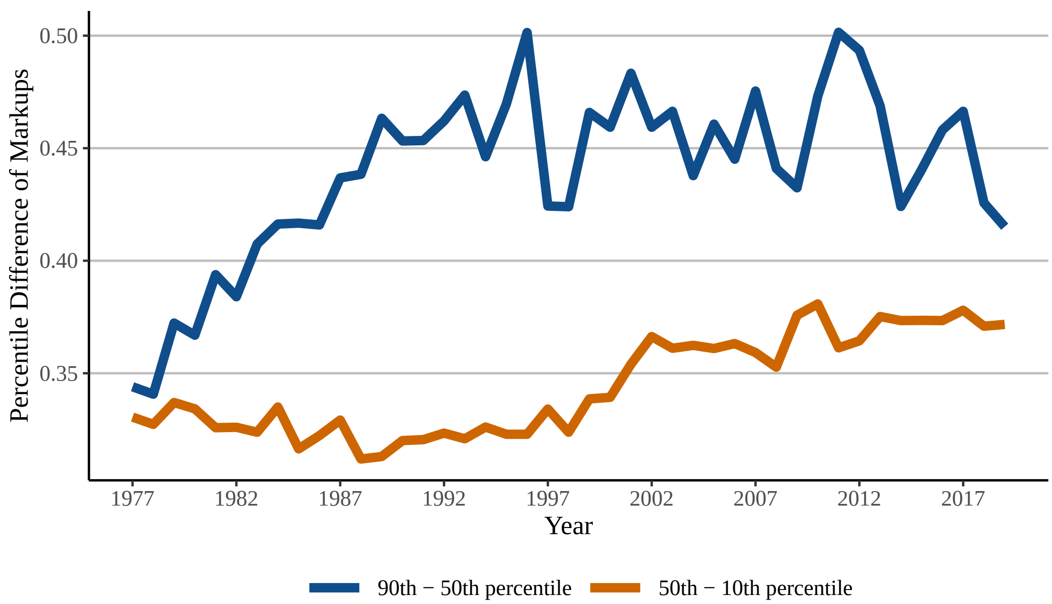
Next, we examine productivity dispersion. Figure 7 reports various measures of productivity dispersion. Productivity dispersion of the sample increases significantly from 1977 to 2019, which is consistent with previously documented evidence (Kehrig, 2015; Andrews, Criscuolo and Gal, 2015, 2016; Barth et al., 2016; Akcigit and Ates, 2021; Gouin-Bonenfant, 2022).[20] In our sample, the standard deviation of log labor productivity almost doubles from 0.52 in 1977 to 1.01 in 2019 (Figure 7a). Moreover, the right-tail of the distribution grows more over this time period than

[20] See Syverson (2011) for a more comprehensive review of productivity dispersion.

Figure 5: Price Markup Cross-Sectional Dispersion



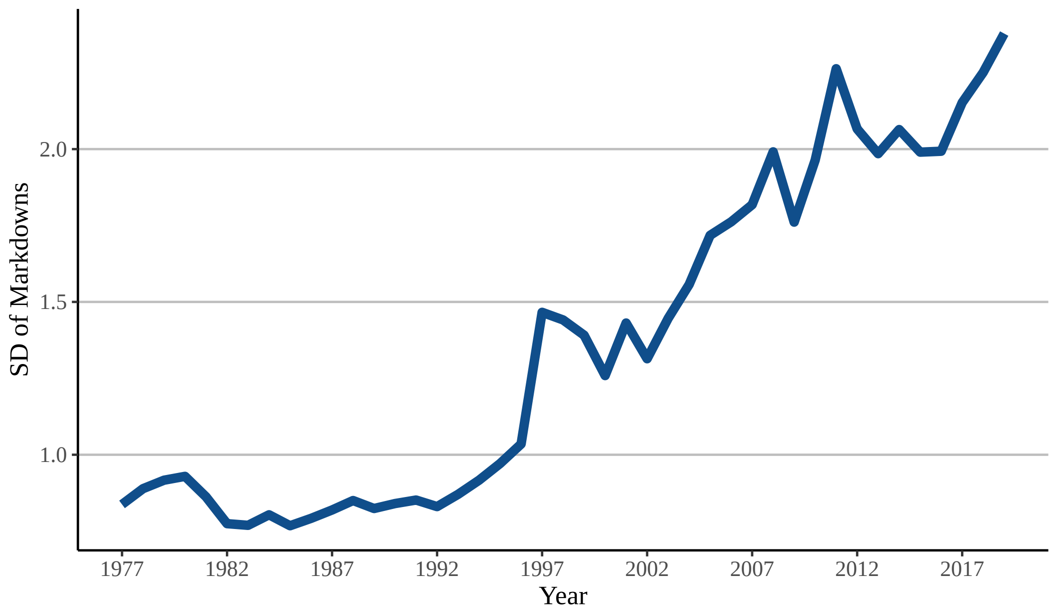
(a) Standard Deviation of Price Markups



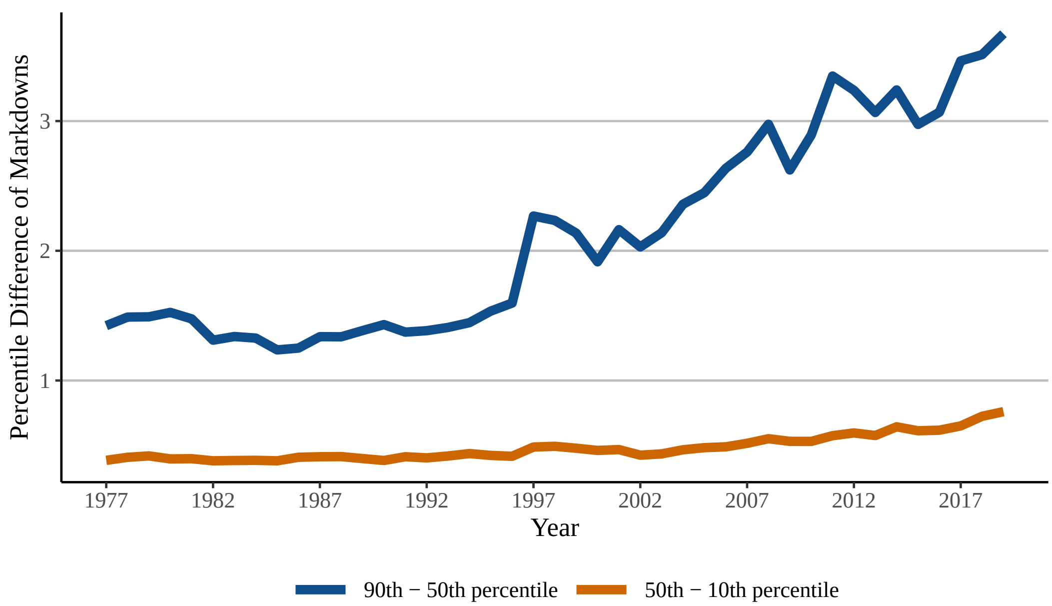
(b) 90/50 and 50/10 Difference of Price Markups

Notes: This figure reports various measures of dispersion of the cross-section of price markups across time. Panel (a) shows the standard deviation of price markups across the years in the sample. Panel (b) reports the difference between the 90th percentile and 50th percentile price markup (blue line) and the difference between the 50th percentile and 10th percentile price markup (orange line) across time. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure 6: Wage Markdown Cross-Sectional Dispersion



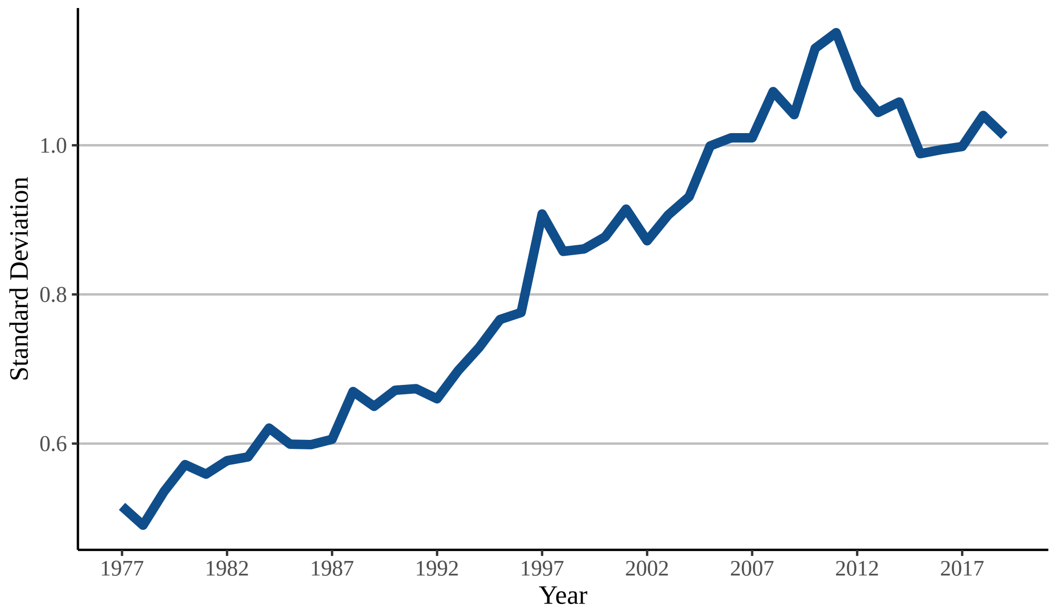
(a) Standard Deviation of Wage Markdowns



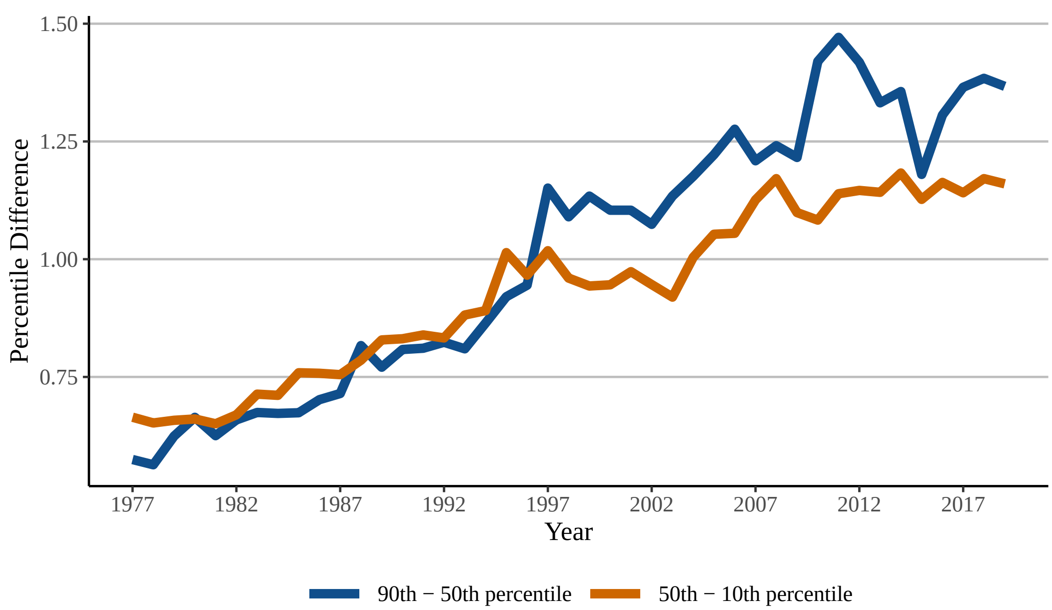
(b) 90/50 and 50/10 Difference of Wage Markdowns

Notes: This figure reports various measures of dispersion of the cross-section of wage markdowns across time. Panel (a) shows the standard deviation of wage markdowns across the years in the sample. Panel (b) reports the difference between the 90th percentile and 50th percentile wage markdown (blue line) and the difference between the 50th percentile and 10th percentile wage markdown (orange line) across time. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure 7: Productivity Cross-Sectional Dispersion



(a) Standard Deviation of Log Labor Productivity



(b) 90/50 and 50/10 Difference of Log Labor Productivity

Notes: This figure reports various measures of dispersion of log labor productivity. Panel (a) shows the standard deviation of log labor productivity across the years in the sample. Panel (b) reports the difference between the 90th percentile and 50th percentile log labor productivity (blue line) and the difference between the 50th percentile and 10th percentile log labor productivity (orange line) across time. All figures are rounded in accordance with U.S. Census disclosure requirements.

the left-tail (Figure 7b). The 90/50 difference more than doubles from 0.58 in 1977 to 1.37 in 2019 whereas the 50/10 difference increases from 0.67 in 1977 to 1.16 in 2019.

We explore the connection between TFP dispersion, our cross-sectional findings, and the aggregate dynamics of markups and markdowns in the theoretical and quantitative analyses presented in Sections 5 and 6. To preview the key intuition, as highly productive firms become even more productive, they gain greater labor market power, leading to an increase in the aggregate wage markdown. However, because of the characteristics of residual product demand, aggregate markups remain relatively stable despite heightened competitive pressure. At the same time, rising productivity dispersion amplifies the cross-sectional variation in both markups and markdowns. Finally, the increase in aggregate labor market power contributes to the secular decline in the labor share and the corresponding rise in the aggregate profit share.

4.3 Firm-Level Cross-Sectional Relationships

What types of firms exhibit high markdowns, and which firms charge high markups? Does the conventional view that larger firms command higher markups hold in the data? To answer these questions, we examine the cross-sectional relationships between firm-level market power measures and key firm characteristics. These insights are critical for understanding misallocation, growth patterns, and the forces driving the aggregate time series trends.[21] The distributions of markup and markdown estimates are very similar regardless of whether we impose the CRS restriction or apply adjustments and all results that follow use these baseline specifications.[22]

Figure 8 reports the relationship between markups, markdowns, and various firm size measures. Due to U.S. Census disclosure requirements, we cannot display scatter plots of markups, markdowns, and firm size measures. However, we are able to report how market power varies across deciles of firm size measures. We specifically regress

$$y_{i,t} = \sum_{j=2}^{10} \beta_j \times \mathbf{1}\left\{i \in D_t^{x,j}\right\} + \theta_{i,t} + \varepsilon_{i,t}, \tag{16}$$

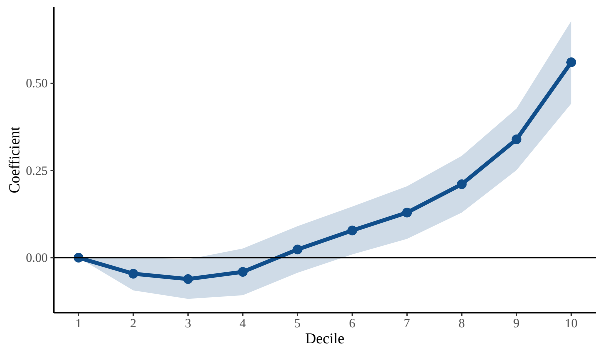
where $D_t^{x,j}$ is the set of observations whose value of characteristic x falls in the j-th decile in year t, $\theta_{i,t}$ is a vector of controls and fixed effects, and $\varepsilon_{i,t}$ is an error term.[23] Figure 8 shows that markdowns are strongly positively related to log value-added labor productivity, while markups exhibit a strong negative relationship. Similarly, markdowns are strongly positively related to sales, whereas markups appear to be negatively related. Lastly, markdowns are strongly negatively related to labor share, whereas markups, if anything, show a positive relationship with labor

[21] Tables A2 to A4 in Appendix A.2 report summary statistics for markups and markdowns in 1977, 2000, and 2019, respectively.

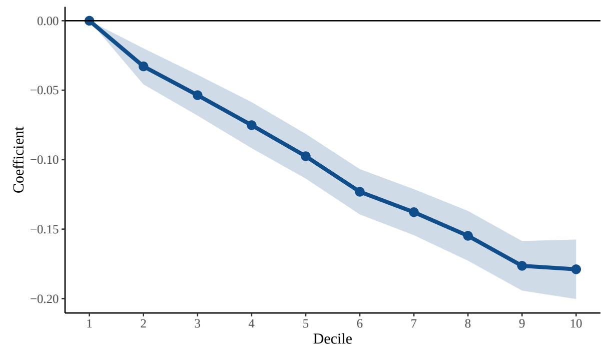
[22] In Tables A5 and A6 we regress estimates that impose the CRS restriction onto unrestricted estimates under various specifications (no fixed effects; NAICS2 × year fixed effects; NAICS2 × year + firm fixed effects). Given the strong correspondence, we adopt the CRS-restricted estimates (and the unadjusted markdowns) as our baseline.

[23] Note that we normalize the first decile to 0 and all coefficients are relative to the first decile.

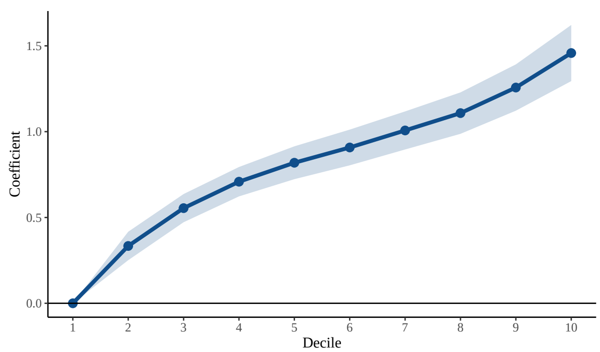
Figure 8: Markup and Markdown Across Deciles of Firm Size Measures



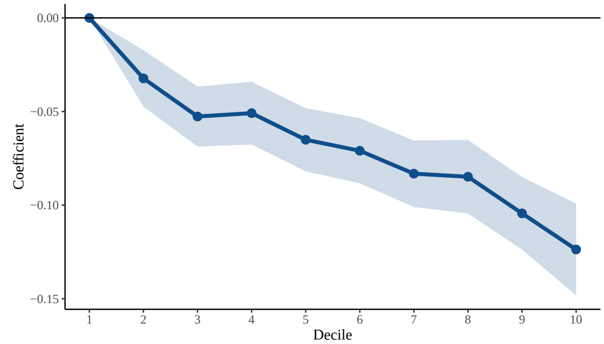
(a) Markdown Across Deciles of Log Labor Prod



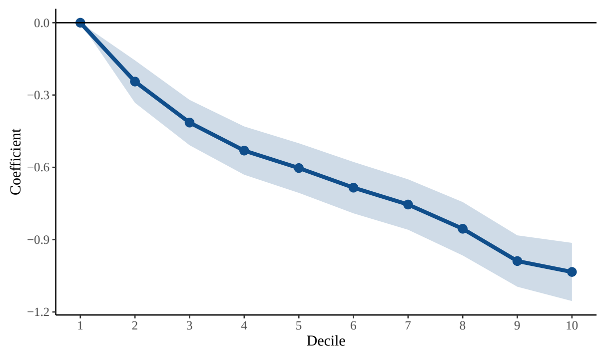
(b) Markup Across Deciles of Log Labor Prod



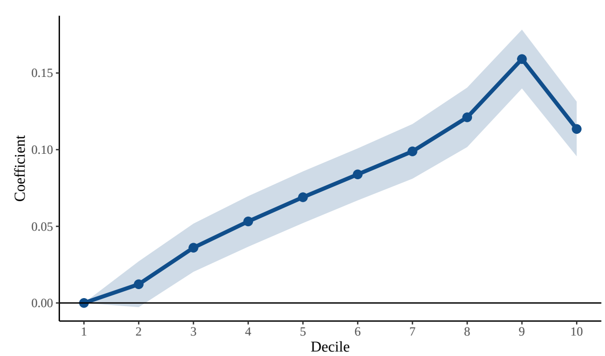
(c) Markdown Across Deciles of Log Sales



(d) Markup Across Deciles of Log Sales



(e) Markdown Across Deciles of Log Labor Share



(f) Markup Across Deciles of Log Labor Share

Notes: This figure reports the relationship between price markups and wage markdowns and various measures of firm size, as well as value-added labor share. Each panel reports the coefficients of the decile dummies from estimating Equation (16) with NAICS2 $\times$ year fixed effects. Panels (a) and (b) present the coefficients for log value-added labor productivity. Panels (c) and (d) display the coefficients for log revenue. Panels (e) and (f) display the coefficients for log value-added labor shares. All figures are rounded in accordance with U.S. Census disclosure requirements.

share.[24] Figure A12 in Appendix A.1 replicates the figures in Figure 8, now including firm fixed effects as a robustness check. The results are broadly consistent.

However, these findings contradict Marshall's Second Law of Demand and a large body of literature on misallocation, which emphasizes a positive correlation between markups and size as a driver of inefficiencies in resource allocation (e.g. Peters, 2020; Baqaee and Farhi, 2020; Edmond, Midrigan and Xu, 2023; Aghion et al., 2023). The negative correlation between markups and size that we uncover suggests that size-dependent taxes or antitrust policies, as implied by Marshall's Second Law of Demand, might not be optimal solely from the perspective of product markups. On the other hand, the positive relationship between size and markdowns might suggest a size-dependent labor-market antitrust policy.

The negative markup-size relationship that we find seemingly contradicts findings on incomplete pass-through of idiosyncratic productivity or cost shocks to prices (e.g. Atkeson and Burstein, 2008; Burstein and Gopinath, 2014; Amiti, Itskhoki and Konings, 2019). Recall that, in economic environments without monopsony power, Marshall's Second Law of Demand and incomplete pass-through are closely relate. A negative productivity or positive cost shock reduces a firm's markup, implying that more productive or larger firms have higher markups. Since marginal costs are not directly observed, the usual empirical strategy is to use productivity or cost shocks of competitive inputs to infer pass-through. The standard interpretation assumes that the firm treats all inputs competitively. However, in the presence of input market power, the results of this empirical approach reflect both pass-throughs arising from markups and markdowns rather than incomplete pass-through of markups alone. We elaborate on this intuition further in Section 5.

Similar to the model and results of Berger, Herkenhoff and Mongey (2022), the positive markdown-size relationship might imply incomplete pass-through of idiosyncratic firm productivity shocks to wages. This is consistent with well-documented empirical findings in labor economics (e.g. Staiger, Spetz and Phibbs, 2010; Card et al., 2018; Kline et al., 2019; Brooks et al., 2021; Chan, Salgado and Xu, 2023). Again, we will analyze the interplay between monopoly and monopsony power, and their connection to pass-throughs, in Section 5.

The size correlations of markups and markdowns also provide insights into the cross-sectional factors underlying the observed aggregate trends. The modest negative correlation between firm markups and size helps explain the relative stability of the aggregate markup over time, despite increasing dispersion and heightened right-skewness in the distribution. Similarly, the positive markdown-size relationship explains the higher wage bill-weighted aggregate markdown compared to the unweighted measure (Figure A9). The unweighted markdown grew more slowly than the wage bill-weighted measure from 1977 to 2002, suggesting that labor reallocation to high-markdown firms was the main driver of markdown growth during that period. However, after 2002, the unweighted markdown increased more quickly, suggesting that within-firm increases in

[24] Despite the more popular prior of a positive firm-size relation, David and Venkateswaran (2019), Díez, Fan and Villegas-Sánchez (2021), and Mertens and Mottironi (2023) find no relationship, or negative relationship between markups and firm size. Moreover, Seegmiller (2023) also finds a positive relationship between markdowns and productivity.

markdown were also a key driver of the aggregate markdown's growth after 2002.

To succinctly report relations of markup and markdown to other firm characteristics, we consider a simple regression specification given by

$$y_{i,t} = \beta x_{i,t} + \theta_{i,t} + \varepsilon_{i,t}, \tag{17}$$

where $y_{i,t}$ is the outcome variable (log markup, log markdown), $x_{i,t}$ is a firm characteristic, $\theta_{i,t}$ is a vector of potential fixed effects, and $\varepsilon_{i,t}$ is an idiosyncratic error term.

The results from estimating Equation (17), using log markups and log markdowns as the respective outcomes, are presented in Tables 3 and 4. We examine firm characteristics such as the markdown/markup, productivity and size metrics, profitability, and labor share. Our main regression specification includes NAICS2 $\times$ year fixed effects since production functions are estimated at the NAICS2 level and to control for industry-level and aggregate trends. Alternative specifications that only have NAICS2 fixed effects or include firm fixed effects are contained in Tables A7 to A10 in Appendix A.2 for robustness. These are results are broadly consistent across specifications and suggest similar dynamics both between and within firms. Markdowns vary significantly more than markups with respect to firm characteristics, which is consistent with previous findings that markdowns exhibit more dispersion than markups. Furthermore, in Column (1) of both Tables 3 and 4 we find that markups and markdowns are negatively related to each other. From both regression results, an increase in the firm's markup is associated with a proportionally greater decrease in the firm's markdown. This result is consistent with the findings of Mertens and Mottironi (2023). We examine the remaining coefficients to better understand the magnitude of this trade-off.

Value-added productivity is our preferred measure of productivity, since TFPR, or revenue-based total factor productivity, is more mechanically related to markups. A simple example is an economy with CES monopolistic competition in the product market and a linear production function: in this case, TFPR is completely uninformative about quantity-based total factor productivity (TFPQ).[25] With the Marshall anti-Second Law of Demand, both markup and TFPR decrease with TFPQ, so the relation between markup and TFPR is positive and the measured covariance of TFPR and markup is upward biased. On the other hand, since valued-added labor productivity also reflects material and capital usage, it better captures variation in TFPQ despite this bias. We present relations of markup and markdown to productivity measures (the firm's TFPR and log labor productivity) and to direct size measures (the firm's sales and wage bill) in Columns (2) to (5)

[25] To see this more clearly, note that

$$\text{TFPR} \;=\; p_i A \;=\; \mu\,\text{MC}\,A \;=\; \mu\,\frac{w}{A}\,A \;=\; \mu w$$

Therefore, TFPR is completely uninformative about TFPQ $= A$, due to complete passthrough. Deviating from the CES demand system, if markup decreases in TFPQ ($\frac{\partial \mu(a)}{\partial a} < 0$), then an increase in TFPQ lowers both markup and TFPR simultaneously, yielding a positive correlation between TFPR and TFPQ, even though TFPQ and markup are negatively related.

Table 3: Price Markups and Firm Characteristics (CRS)

	Log Price Markup						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Wage Markdown	-0.105						
	(0.005)						
Log TFPR		0.000					
		(0.016)					
Log Labor Productivity			-0.051				
			(0.005)				
Log Sales				-0.017			
				(0.001)			
Log Wage Bill					-0.001		
					(0.001)		
Profit Share						-0.019	
						(0.018)	
Log Labor Share VA							0.045
							(0.005)
NAICS2 FE	No	No	No	No	No	No	No
NAICS2 $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	No	No	No
Observations	69,500	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log price markups onto firm characteristics with NAICS2 $\times$ year fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Column (1) examines log wage markdowns, Columns (2) and (3) analyze log TFPR and log labor productivity, Columns (4) and (5) assess log sales and log wage bill, and Columns (6) and (7) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

of Tables 3 and 4.[26] Similarly to Figure 8, the results indicate that markups have no association with size or productivity (point estimates from 0 to -0.05 and all statistically significant at the usual levels), whereas markdowns are positively related to a firm's productivity and size (point estimates from 0.13 to 1.19).

Finally, we examine the associations between firm profitability and labor share with markups and markdowns (Columns (6) and (7), respectively, in Tables 3 and 4).[27] Markups exhibit a modest negative relationship with profitability and a modest positive relationship with the labor share, whereas markdowns have the opposite signs with coefficients of larger magnitude. The results are counterintuitive for markups as usually a higher markup implies a higher level of profitability

[26]We follow the definition and variable construction of labor productivity in Donangelo et al. (2019); Appendix B.2 provides more information on its construction.

[27]Appendix B.2 provides more information on how firm profitability and labor share of value-added are constructed and defined.

Table 4: Wage Markdowns and Firm Characteristics (CRS)

	Log Wage Markdown						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Price Markup	-2.599						
	(0.056)						
Log TFPR		1.188					
		(0.135)					
Log Labor Productivity			0.191				
			(0.016)				
Log Sales				0.203			
				(0.009)			
Log Wage Bill					0.129		
					(0.007)		
Profit Share						0.568	
						(0.095)	
Log Labor Share VA							-0.385
							(0.022)
NAICS2 FE	No	No	No	No	No	No	No
NAICS2 $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	No	No	No
Observations	69,500	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log wage markdowns onto firm characteristics with NAICS2 $\times$ year fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Column (1) examines log price markups, Columns (2) and (3) analyze log TFPR and log labor productivity, Columns (4) and (5) assess log sales and log wage bill, and Columns (6) and (7) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

and lower labor share. However, these are univariate regressions and markup and markdowns are negatively related. The previous results suggest that firms are able to increase markdowns at a rate that more than offsets the proportional decrease in markups. Thus, the specifications in Columns (6) and (7) in Table 3 do not account for markdowns changing and thus lead to the counterintuitive result. The corresponding results for markdowns are more straightforward, a 10 percentage point increase in profitability is associated with a 0.04 log point increase in markdowns and 1% increase in the labor share is associated with a -0.39% decrease in markdowns. Taken together, these results are also consistent with the trends of a declining aggregate labor share and rising corporate profitability.

However, to reconcile our finding with the previous literature, we also consider a specification in which we regress the log of the product of markups and markdowns, i.e., $\ln \nu_{i,t}\mu_{i,t}$. We call this the total wedge, and it (imperfectly) approximates the total market power of the firm. We regress

Table 5: Total Wedge and Firm Characteristics (CRS)

	Log Total Wedge					
	(1)	(2)	(3)	(4)	(5)	(6)
Log TFPR	1.189					
	(0.124)					
Log Labor Productivity		0.140				
		(0.016)				
Log Sales			0.186			
			(0.009)			
Log Wage Bill				0.128		
				(0.007)		
Profit Share					0.549	
					(0.092)	
Log Labor Share VA						-0.340
						(0.020)
NAICS2 FE	No	No	No	No	No	No
NAICS2 $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	No	No	No	No
Observations	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log total wedges onto firm characteristics with NAICS2 $\times$ year fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Columns (1) and (2) analyze log TFPR and log labor productivity, Columns (3) and (4) assess log sales and log wage bill, and Columns (5) and (6) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

the total wedge onto the same set of productivity, size, and factor share measures. Table 5 reports the coefficients from the main specification that includes NAICS2 $\times$ year fixed effects. Tables A11 and A12 in Appendix A.2 report the results with only NAICS2 fixed effects and including firm fixed effects, respectively. The firm's total wedge increases with productivity and size, indicating that the firm's total market power, as measured by the total wedge of labor, rises alongside firm productivity and size. Furthermore, the firm's profitability and labor share are positively and negatively related, respectively, to the total wedge.

These results suggest that larger firms exhibit greater total market power and are underproducing relative to smaller firms, contributing to market power-based misallocation. However, understanding how the division between markups and markdowns affects this misallocation requires a more detailed quantitative analysis. Additionally, the positive relationship between profitability and the total wedge, alongside the negative relationship between labor share and the total wedge, is consistent with the notion that firms with greater market power are more profitable and have smaller labor shares. Similar to the results on productivity and size, we return to this evidence when discussing the quantitative model and the implications of the division of market power.

5 Model

We develop a general yet parsimonious framework in which firms exercise both monopoly power in the product market and monopsony power in the labor market to rationalize and quantify the empirical patterns in Section 4. Output and labor each enter through a Homothetic Single Aggregator (HSA) system, following Matsuyama and Ushchev (2017) and Matsuyama (2023), which imposes only monotonicity and curvature conditions on residual demand while retaining analytical tractability.[28] In principle, the most general demand system would require the full distribution of prices to capture the competitive environment, which is analogous to the Slutsky matrix in a finite-goods setting. However, HSA reduces the dimensionality of relevant prices to a single index while preserving homotheticity. Notably, the constant elasticity of substitution (CES) and translog demand systems emerge as special cases of HSA.

We extend HSA to residual labor supply, allowing for a unified analysis of how markups and markdowns respond to firm- and market-level shocks. In this section we impose no parametric functional forms; later, we nonparametrically recover product-demand and labor-supply system from our cross-sectional estimates. By contrast, a CES aggregator, though analytically convenient, imposes constant elasticities of substitution and hence constant markups and markdowns. HSA, by letting elasticities vary across firms and markets, captures the rich size- and productivity-dependence of markups and markdowns documented in Section 4.[29]

The remainder of this section lays out the model environment and agents (Sections 5.1 to 5.3), characterizes its equilibrium (Section 5.4), analyzes its key properties (Section 5.5), and discusses the calibration (Section 5.6). All omitted proofs appear in Appendix E.

5.1 Environment

The economy consists of a representative household, a continuum of heterogeneous intermediate firms, and a competitive final good firm. Time is discrete. There is a unit measure of atomistic households that supply labor to firms and make the usual consumption and investment decision. The households also own the firms and are rebated firms' profits. There is a unit mass of intermediate firms that are indexed by $i \in [0, 1]$. These firms draw their productivity $\omega_i \in (0, \infty)$

[28] Recent applications of HSA include Matsuyama and Ushchev (2021*a*,*b*, 2022, 2023*a*,*b*), Wang and Werning (2022), Grossman, Helpman and Lhuillier (2023), and Baqaee, Farhi and Sangani (2024*b*).

[29] A double-nested CES with oligopolistic product and oligopsonistic labor competition (e.g. extending and combining the models of Atkeson and Burstein (2008) and Berger, Herkenhoff and Mongey (2022)) can generate size-markup dynamics but remains relatively more restrictive. For example, markups in one industry do not respond to competitive changes in another. To retain block recursivity, it also requires firms to compete with the same set of firms in both labor and product markets.

from an invariant distribution $F(\omega)$ at $t = 0$. Note that here we denote ω_i as the level of firm productivity rather than the log. We assume that once firms draw their productivity it is fixed and that there is no entry or exit. In other words, there is a fixed firm type distribution with no entry and exit dynamics. Firms are local monopolists in their output market and local monopsonists in their labor market. Firms also take aggregates as given.

Intermediate firms produce tradable differentiated goods $y_{i,t}$ and operate a value-added production function that uses only labor $l_{i,t}$ as an input.[30] The production function is given by

$$y_{i,t} = \omega_i l_{i,t}. \tag{18}$$

The final goods firm competitively bundles the differentiated goods $y_{i,t}$ according to the following technology

$$\int_{i=0}^{1} r_i \left(\frac{p_{i,t}}{P(\mathbf{p}_t)} \right) di = 1, \tag{19}$$

where $r_i(\cdot)$ is the revenue share function for intermediate firm i, $p_{i,t}$ is the price of firm i's output, $\mathbf{p}_t$ is the vector of all firms' prices, and $P(\mathbf{p}_t)$ is the price aggregator (let $P_t \equiv P(\mathbf{p}_t)$) that determines the competition price index. The competition price index P_t is implicitly defined by Equation (19) and we have

$$r_i \left(\frac{p_{i,t}}{P_t} \right) = \frac{p_{i,t} y_{i,t}}{P_t^I Y_t}, \tag{20}$$

where Y_t is aggregate output and P_t^I is the ideal price index, which we normalize to 1 each period.[31] For ease of notation we define $a_{i,t} \equiv p_{i,t}/P_t$.

5.2 Households

Households solve the following problem

$$U_0 = \max_{\{l_{i,t}, C_t, A_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(C_t, L_t),$$
$$\text{subject to} \tag{21}$$
$$C_t + A_{t+1} = \int_{i=0}^{1} w_{i,t} l_{i,t}\, di + R_t A_t + \Pi_t.$$

where $\beta \in (0, 1)$ is the subjective time discount factor, C_t is aggregate consumption, L_t is aggregate labor, A_t is the total holdings in the risk-free asset from period $t-1$, Π_t is aggregate profit, $w_{i,t}$ is the

[30] We consider only labor for simplicity; adding physical capital or intangible capital do not change the core results.

[31] In general $P_t \neq P_t^I$; when the aggregator is CES, $P_t = cP_t^I$ for some scalar $c \in \mathbb{R}_{++}$. The competition price index P_t is used by firms to determine their revenue share of the total economy and reflects the level of product market competition in the economy. The competition price P_t also reflects all the cross-price effects in the demand system. The ideal price index P_t^I is the final price of the aggregated final good for consumers and is used for welfare calculations. Given P_t and $r_i(\cdot)$, P_t^I can be implicitly defined by Equation (20). We formally define P_t^I and derive its relationship with P_t in Appendix E.1. See also Matsuyama and Ushchev (2017) for a more comprehensive derivation and proof.

wage from firm i, $l_{i,t}$ is the labor supplied to firm i, and R_t is the gross risk-free rate. The risk-free asset has a net-zero supply. The aggregate supply labor index and consequently the aggregate wage index is given by

$$\int_{i=0}^{1} s_i \left(\frac{w_{i,t}}{W(\mathbf{w}_t)} \right) di = 1, \tag{22}$$

and we have that

$$s_i \left(\frac{w_{i,t}}{W_t} \right) = \frac{w_{i,t} l_{i,t}}{W_t^I L_t}, \tag{23}$$

where $s_i(\cdot)$ is the wage bill share function of firm i, $\mathbf{w}_t$ is the vector of all firms' wages, $W(\mathbf{w}_t)$ is the aggregator (similarly define $W_t \equiv W(\mathbf{w}_t)$) for the competition wage index, and W_t^I is the ideal wage index.[32] The household's period utility function $u(\cdot,\cdot)$ satisfies the usual regularity conditions. Similar to before we define $b_{i,t} \equiv w_{i,t}/W_t$.

5.3 Firms

The intermediate firms solve the following

$$\begin{gathered} \max_{p_{i,t}, w_{i,t}} \pi_{i,t} = p_{i,t} y_{i,t} - w_{i,t} l_{i,t}, \\ \text{subject to} \\ p_{i,t} y_{i,t} = r_i\left(a_{i,t}\right) P_t^I Y_t, \\ w_{i,t} l_{i,t} = s_i\left(b_{i,t}\right) W_t^I L_t, \\ y_{i,t} = \omega_i l_{i,t}, \end{gathered} \tag{24}$$

Notice that there is an equivalent formulation of this problem in which the firm picks labor and output rather than prices. However, given the form of the aggregators, the formulation in which firms choose prices is generally more straightforward to solve. Before we inspect firm's optimality conditions, it is useful to define other quantities. First, we can relate the market shares with the elasticity of residual product demand and labor supply as follows, respectively,

$$\varepsilon_i\left(a_{i,t}\right) \equiv -\frac{\partial \ln y_{i,t}}{\partial \ln p_{i,t}} = 1 - \frac{r_i'\left(a_{i,t}\right) a_{i,t}}{r_i\left(a_{i,t}\right)}, \tag{25}$$

$$\eta_i\left(b_{i,t}\right) \equiv \frac{\partial \ln l_{i,t}}{\partial \ln w_{i,t}} = \frac{s_i'\left(b_{i,t}\right) b_{i,t}}{s_i\left(b_{i,t}\right)} - 1. \tag{26}$$

[32]The same intuition from the product side also applies here. Firms take W_t to determine their wage bill share whereas consumers take W_t^I as the aggregate wage received and is used for welfare calculations. Similar to the price indices, the wage indices are implicitly defined by Equations (22) and (23). We also formally define W_t^I and its relationship with W_t in Appendix E.1.

Given the definition of price markups and wage markdowns in terms of their respective elasticities it also follows that

$$\mu_i\left(a_{i,t}\right) \equiv \frac{\varepsilon_i\left(a_{i,t}\right)}{\varepsilon_i\left(a_{i,t}\right)-1} = \frac{r_i\left(a_{i,t}\right)-r_i'\left(a_{i,t}\right)a_{i,t}}{r_i'\left(a_{i,t}\right)a_{i,t}}, \tag{27}$$

$$\nu_i\left(b_{i,t}\right) \equiv \frac{\eta_i\left(b_{i,t}\right)+1}{\eta_i\left(b_{i,t}\right)} = \frac{s_i'\left(b_{i,t}\right)b_{i,t}}{s_i'\left(b_{i,t}\right)b_{i,t}-s_i\left(b_{i,t}\right)}. \tag{28}$$

Complete derivations are in Appendix E.2. Notably, under the HSA aggregator, each firm's markup and markdown depend only on $a_{i,t}$ and $b_{i,t}$, respectively. Hence, aggregate conditions and the competitive positions of other firms enter each firm's price and wage setting process in a parsimonious way.

Combining the firm's first-order conditions with Equations (27) and (28) yields the pricing and wage rules, which are given by

$$p_{i,t} = \mu_i\left(a_{i,t}\right)\nu_i\left(b_{i,t}\right)\frac{w_{i,t}}{\omega_i}, \tag{29}$$

$$w_{i,t} = \nu_i\left(b_{i,t}\right)^{-1}\mu_i\left(a_{i,t}\right)^{-1}p_{i,t}\omega_i. \tag{30}$$

Notice that the marginal cost term $\nu_i\left(b_{i,t}\right)w_{i,t}\omega_i^{-1}$ and marginal revenue product of labor term $\mu_i\left(a_{i,t}\right)^{-1}p_{i,t}\omega_i$ include the markdown and markup, respectively. This implies that the firm's price and wage setting behavior is influenced by both its product demand and labor supply as shown by Kroft et al. (2023) and Trottner (2023).

5.4 Equilibrium

Before we define the equilibrium, we discuss the assumptions that we must impose onto $r_i(\cdot)$ and $s_i(\cdot)$ for the model to ensure that proper preferences are defined. We also discuss any additional assumptions that we need to ensure the existence and uniqueness of the equilibrium. The assumptions are as follows:

Assumption 6 (Smoothness). *Both $r_i(\cdot)$ and $s_i(\cdot)$ are at least twice continuously differentiable for all $i \in [0,1]$.*

Assumption 7 (Basic Monotonicity Conditions). *$r_i'\left(a_{i,t}\right) < 0$ and $s_i'\left(b_{i,t}\right) > 0$ for all $i \in [0,1]$.*

Assumption 8 (First Law of Demand and Supply). *$r_i(\cdot)$ and $s_i(\cdot)$ satisfy the following conditions*

$$r_i'\left(a_{i,t}\right)a_{i,t} < r_i\left(a_{i,t}\right), \quad s_i'\left(b_{i,t}\right)b_{i,t} > s_i\left(b_{i,t}\right),$$

for all $i \in [0,1]$.

Assumption 9 (Strict Monotonicity Conditions). *$a_{i,t}\mu_i\left(a_{i,t}\right)^{-1}$ is strictly increasing in $a_{i,t}$ and $b_{i,t}\nu_i(b_{i,t})$ is strictly increasing in $b_{i,t}$ for all $i \in [0,1]$. That is*

$$\frac{\partial a_{i,t}\mu_i\left(a_{i,t}\right)^{-1}}{\partial a_{i,t}} > 0 \quad \frac{\partial b_{i,t}\nu_i\left(b_{i,t}\right)}{\partial b_{i,t}} > 0.$$

Assumptions 6 to 8 are standard. These assumptions are especially needed to guarantee a well-defined equilibrium in a monopolistic and monopsonistic competition environment. In particular, the combination of Assumptions 7 and 8 ensure well-defined elasticities and markups/markdowns. Assumption 9 ensures that the equilibrium is unique by imposing an additional monotonicity condition. Definition 1 defines the equilibrium.

Definition 1 (Equilibrium). *An equilibrium consists of a set of prices $\mathbf{p}_t$, wages $\mathbf{w}_t$, and allocations $\{y_{i,t}, l_{i,t}\}_{i\in[0,1]}$ for all $t \geq 0$ such that:*

1. *It solves the household's problem (21) and the firm's problem (24) for all $i \in [0,1]$.*

2. *All markets clear every period.*

5.5 Model Properties

Before discussing the calibration and model estimation strategy, we examine the model's properties theoretically. We show that the model, under the specified conditions, can qualitatively match the empirical findings from Section 4. We also discuss contexts and quantities that depend not only on total market power but also the decomposition between product and labor market power. Furthermore, we provide a more detailed discussion on how the model reconciles our findings with previous evidence of incomplete pass-throughs.

If the model satisfies Assumptions 6 to 9 as well as Assumption 10, which is defined below, then it is able to qualitatively replicate several of cross-sectional findings in Section 4.3. Only Assumption 10 for demand represents a key departure in terms of assumptions, as most prior work assume Marshall's Second Law of Demand.[33] Proposition 3 formally presents this result.

Assumption 10 (Anti-Second Law of Demand and Second Law of Supply). *The elasticity of product demand and labor supply are strictly decreasing in $a_{i,t}$ and $b_{i,t}$, respectively, for all $i \in [0,1]$. That is,*

$$\frac{\partial \varepsilon_i\left(a_{i,t}\right)}{\partial a_{i,t}} < 0, \quad \frac{\partial \eta_i\left(b_{i,t}\right)}{\partial b_{i,t}} < 0.$$

Proposition 3 (Comparative Statics of Firm Characteristics). *Suppose the conditions of Assumptions 6 to 10 hold, then holding aggregate conditions constant, an increase in firm i's productivity ω_i leads to a*

[33]Marshall's Second Law of Demand formally requires that $\frac{\partial \varepsilon_i\left(a_{i,t}\right)}{\partial a_{i,t}} > 0$.

decrease in its price $p_{i,t}$ and price markup $\mu_i\left(a_{i,t}\right)$, while its revenues $p_{i,t}y_{i,t}$, wage $w_{i,t}$, wage bill $w_{i,t}l_{i,t}$, and wage markdown $\nu_i\left(b_{i,t}\right)$ increase.

Proof. See Appendix E.3 for the proof. □

Note that Proposition 3 does not address how firm labor shares and profit shares vary with productivity. The firm-level labor share is given by

$$\frac{w_{i,t}l_{i,t}}{p_{i,t}y_{i,t}} = \frac{1}{\mu_i\left(a_{i,t}\right)\nu_i\left(b_{i,t}\right)}, \tag{31}$$

which follows directly from combining the pricing rule (29) with the production function (18). Since the model generates the observed negative relationship between markups and markdowns, Equation (31) reflects the trade-off in how firms extract rents as productivity varies. While the level of profits increases with productivity in line with the logic of profit maximization, Assumptions 7 to 10 alone do not determine how labor and profit shares change with respect to productivity. Therefore, determining how labor and profit shares vary with productivity is inherently a quantitative question. We address this in Section 6.

What does our model predict when the firm faces a TFP or cost shock? We define the price pass-through $\rho^p_{i,t}$ and wage pass-through $\rho^w_{i,t}$ as follows:

$$\rho^p_{i,t} \equiv \frac{\partial \ln p_{i,t}}{\partial \ln \mathrm{MC}_{i,t}}, \tag{32}$$

$$\rho^w_{i,t} \equiv \frac{\partial \ln w_{i,t}}{\partial \ln \mathrm{MRPL}_{i,t}}. \tag{33}$$

We refer to these as the primitive price and wage pass-throughs, respectively. However, empirically, marginal costs and marginal revenue products are not directly observed. As a result, many studies either utilize a proxy variable or are explicitly interested in pass-throughs with respect to other variables. A common candidate is the firm's productivity or some closely related proxy (e.g. labor productivity). We define the pass-throughs with respect with productivity as follows

$$\rho^{p,\omega}_{i,t} \equiv \frac{\partial \ln a_{i,t}}{\partial \ln \omega_i}, \tag{34}$$

$$\rho^{w,\omega}_{i,t} \equiv \frac{\partial \ln b_{i,t}}{\partial \ln \omega_i}. \tag{35}$$

We call Equations (34) and (35) the effective price and wage pass-throughs, respectively.[34] We express the pass-throughs in terms of their relative price/wage to the competitive index for ease of exposition. Most empirical studies estimating pass-throughs estimate some variant of Equations (34) and (35). We can relate Equation (32) to Equation (34) and Equation (33) to Equation (35) through

[34]We use the terms effective pass-through and measured pass-through interchangeably throughout the text.

the chain rule. This results in the following

$$\rho_{i,t}^{p,\omega} = \frac{\partial \ln a_{i,t}}{\partial \ln \mathrm{MC}_{i,t}} \frac{\partial \ln \mathrm{MC}_{i,t}}{\partial \ln \omega_i} = \rho_{i,t}^{p} \frac{\partial \ln \mathrm{MC}_{i,t}}{\partial \ln \omega_i}, \tag{36}$$

$$\rho_{i,t}^{w,\omega} = \frac{\partial \ln b_{i,t}}{\partial \ln \mathrm{MRPL}_{i,t}} \frac{\partial \ln \mathrm{MRPL}_{i,t}}{\partial \ln \omega_i} = \rho_{i,t}^{w} \frac{\partial \ln \mathrm{MRPL}_{i,t}}{\partial \ln \omega_i}. \tag{37}$$

With these definitions and relations, we proceed to Proposition 4, which formally establishes a result that shows how the model reconciles the empirical findings with prior research.

Proposition 4 (Reconciliation of Pass-Throughs)**.** *Suppose Assumptions 6 to 9 hold. Then, the effective price and wage pass-throughs can be expressed as*

$$\rho_{i,t}^{p,\omega} = -\frac{\rho_{i,t}^{p} + \rho_{i,t}^{p}\rho_{i,t}^{w}\eta_i\left(b_{i,t}\right)}{\rho_{i,t}^{p}\varepsilon_i\left(a_{i,t}\right) + \rho_{i,t}^{w}\eta_i\left(b_{i,t}\right)}, \tag{38}$$

$$\rho_{i,t}^{w,\omega} = \frac{\rho_{i,t}^{w}\rho_{i,t}^{p}\varepsilon_i\left(a_{i,t}\right) - \rho_{i,t}^{w}}{\rho_{i,t}^{p}\varepsilon_i\left(a_{i,t}\right) + \rho_{i,t}^{w}\eta_i\left(b_{i,t}\right)}. \tag{39}$$

Proof. See Appendix E.4 for the proof. □

Given Equations (36) and (38), we deduce that

$$\frac{\partial \ln \mathrm{MC}_{i,t}}{\partial \ln \omega_i} = -\frac{1 + \rho_{i,t}^{w}\eta_i\left(b_{i,t}\right)}{\rho_{i,t}^{p}\varepsilon_i\left(a_{i,t}\right) + \rho_{i,t}^{w}\eta_i\left(b_{i,t}\right)}. \tag{40}$$

Under perfectly competitive factor markets and constant returns to scale, Equation (40) simplifies to -1, and the primitive price pass-through and the effective (or measured) price pass-through are equal up to a sign: $\rho_{i,t}^{p} = -\rho_{i,t}^{p,\omega}$.[35] This equivalence is often implicitly assumed in empirical studies when interpreting their estimates of pass-through rates. However, when factor markets are imperfectly competitive, the magnitude of (40) can fall below 1, allowing for $|\rho_{i,t}^{p}| > 1 > |\rho_{i,t}^{p,\omega}|$, thereby reconciling our empirical findings with those of past research.[36] The intuition behind this result lies in the interaction of imperfect competition across both output and factor markets. Under imperfect competition in factor markets, a firm's marginal cost depends not only directly on productivity but also indirectly through the effects of productivity on the markdown and factor prices.

Building on this intuition, we explore how productivity shocks propagate through equilibrium prices via supply-side and demand-side channels. The term $\rho_{i,t}^{w}\eta_i\left(b_{i,t}\right)$ captures the total effect of a change in the firm's marginal revenue product of labor (MRPL) on employment, while $\rho_{i,t}^{p}\varepsilon_i\left(a_{i,t}\right)$

[35]As factor markets approach perfect competition, $\eta_i\left(b_{i,t}\right) \to \infty$ and $\rho_{i,t}^{w} \to 1$, leading to $\frac{\partial \ln \mathrm{MC}_{i,t}}{\partial \ln \omega_i} \to -1$. This result is consistent with the standard assumption of perfectly competitive factor markets.

[36]If Proposition 3 holds, then $\rho_{i,t}^{p} > 1$ and $\varepsilon_i\left(a_{i,t}\right) > 1$. Therefore, in our setting, the magnitude of Equation (40) is always less than 1. See the proof for Proposition 3 in Appendix E.3 and Lemma 5 in Appendix E.4 for details.

captures the total effect of a change in the firm's marginal cost on output.[37] The way productivity shocks affect equilibrium prices depends on how elasticities map into pass-throughs along both input and output margins. On the supply side, input elasticities and pass-throughs determine how markdowns and input prices respond to productivity. On the demand side, the demand elasticity and pass-through determine how markups and output prices adjust. Thus, the relative strength of these supply-side and demand-side forces determines whether productivity shocks amplify or dampen the overall price response. Consequently, the magnitudes of $\rho_{i,t}^{p}$ and $\rho_{i,t}^{p,\omega}$ are determined quantitatively by the degree and nature of competition in factor and product markets. A similar logic applies to wage pass-throughs, as shown in Equations (37) and (39). However, as shown in Section 6, our calibration indicates that although $\rho_{i,t}^{w}$ and $\rho_{i,t}^{w,\omega}$ differ in magnitude, both remain within the interval $(0, 1)$.

Proposition 4 also shows that the effective pass-throughs depend on the decomposition of firm-level market power. Consider an economy with product market power but perfectly competitive labor markets. Firms in both economies are indexed by their productivity in ascending order. Let $\tilde{\mu}_{i,t}$ denote the price markups in this new economy. In equilibrium, $\tilde{\mu}_{i,t} = \mu_{i,t}\nu_{i,t}$, meaning that price markups in the new economy are equivalent to the total wedge in the current economy across all firms. Aside from this distinction, the new economy is otherwise identical to the currently specified economy. Let this new economy be called a markup-equivalent economy.

Following, the same notation convention, the magnitude of the effective price pass-through in the markup-equivalent economy is equal to the magnitude of the price pass-through. That is, $|\tilde{\rho}_{i,t}^{p}| = |\tilde{\rho}_{i,t}^{p,\omega}|$. This follows from our previous result. This implies that the pass-throughs in the different economies are generally different, which results in different comparative statics as firm-level productivity changes.[38] Therefore, the effects of policy and the degree of inefficiency in these two economies are generally different despite having observationally similar distributions of total market power. This implies that decomposing total market power and allocating it correctly is crucial to understand various dynamics.

5.6 Non-parametric Calibration

We now outline the calibration strategy for the model. The primary objective of this calibration is threefold: first, to analyze how TFP dispersion influences the trends in aggregate markups and markdowns; second, to disentangle the channels through which changes in the competitive environment drive these aggregate outcomes. Leveraging this decomposition, we examine the model's implications for reallocation and pass-through dynamics over the sample period. Finally, we connect our findings to the existing literature, particularly regarding pass-through and its

[37] In the special case of a CES labor supply function, $\rho_{i,t}^{w} = 1$ and $\eta_i\left(b_{i,t}\right) > 0$. Here, productivity affects wages but not the markdown. A similar result holds under CES product demand. In general, however, $\rho_{i,t}^{p} \neq -\rho_{i,t}^{p,\omega}$.

[38] By construction, the markup-equivalent economy also cannot reproduce any of the wage pass-throughs that are found in labor economics.

relevance to Marshall's Second Law of Demand. To cleanly isolate the role of TFP dispersion, we non-parametrically identify the demand and supply systems based on data at the beginning of the sample in 1977, but let TFP distribution evolves based on the observed TFP distribution.

We begin by describing the identification of $r_i(\cdot)$ and $s_i(\cdot)$ from the data. The general idea is that we leverage our markup and markdown estimates, together with their joint distribution of revenue and wage bills share to recover $r_i(\cdot)$ and $s_i(\cdot)$. To isolate the effect of TFP dispersion from changes in demand system, we use the joint distribution of markups, markdowns, sales, and wage bills at the beginning of the sample in 1977. We then recover the implied pass-through and productivity, and by integration, we will be able to recover relative prices and wages. We will then fit the recovered relative prices and wages over the observed distribution of $r_i(\cdot)$ and $s_i(\cdot)$ to recover the revenue and wage bills share function over relative prices and wages. Appendix F provides more details on the computational and numerical procedure. Assumption 11 describes the key assumption we use to identify $r(\cdot)$ and $s(\cdot)$ from the data.

Assumption 11 (Demand and Supply Shifters for Identification)**.** *Firms of type i have their idiosyncratic product demand and labor supply system in the following form:*

$$r_i(\cdot) = B_i r\left(\frac{p_i}{P}\right), \quad s_i(\cdot) = C_i s\left(\frac{w_i}{W}\right)$$

With B_i being product's quality shifter and C_i being labor's amenity value for firm type i.

We allow for an unobserved quality shifter for product demand and amenity value for labor supply in an multiplicative manner. A higher quality shifter B_i implies a higher product demand for type i firm given relative price. Similarly, a higher amenity value C_i means higher labor supply given relative wage. One implication of this assumption is that TFP ω_i is not separately identified from B_i and C_i. However, $A_i = B_i C_i \omega_i$ is identified. That is, the quality and amenity shifters behave similarly to productivity, since firms with higher quality and amenity are able to attract more demand and supply without changing their price or wage. Henceforth, we will refer A_i as quality adjusted TFP, or simply TFP.

If Assumption 11 holds, revenue and wage bill shares are strictly increasing in quality adjusted TFP. Therefore, we can identify firms' type by ranking them over their revenue share. Concretely, firms' type are identify based on their position on the revenue cumulative distribution.[39] However, we ultimately want the demand and supply systems to have relative prices and wages as their argument. The following proposition implies that we can use the relation between revenue share and markup (similarly for markdown) to back out effective and primitive pass-through. Once effective and primitive pass-through are identified, we can use the relations between pass-through and markup to back out quality adjusted TFP. Once TFP is identified, relative prices and wages can be obtained by integration, up to integration constants.

[39] Appendix F provides more detail on this procedure.

Proposition 5 (Recovering Pass-Through and Productivity). *Given $\nu_i, \mu_i, r_i(\cdot), s_i(\cdot)$, the following relations identify productivity and pass-through:*

$$\frac{\partial \ln \mu_i}{\partial i} = (1-\mu_i)\left(1-\frac{1}{\rho_i^p}\right)\frac{\partial \ln r_i}{\partial i} = \left(1-\frac{1}{\rho_i^p}\right)\rho_i^{p,\omega}\frac{\partial \ln A_i}{\partial i},$$
$$\frac{\partial \ln \nu_i}{\partial i} = \left(\frac{1}{\nu_i}-1\right)\left(\frac{1}{\rho_i^w}-1\right)\frac{\partial \ln s_i}{\partial i} = \left(\frac{1}{\rho_i^w}-1\right)\rho_i^{w,\omega}\frac{\partial \ln A_i}{\partial i}.$$

Proof. See Appendix F.1 for the proof. □

Proposition 5 establishes that by examining changes in markups (markdowns) relative to revenue (wage bill) shares, we can identify primitive pass-through. Once primitive pass-through is determined, Proposition 4 allows us to identify productivity A_i, which in turns allows us to use the results in Proposition 5 to recover the distributions. Finally, with productivity and effective pass-through identified, we can integrate to recover relative prices and wages. Fitting revenue and wage bill share over relative prices and wages, we can recover $r(\cdot)$ and $s(\cdot)$.

Once $r(\cdot)$ and $s(\cdot)$ are identified, the remainder of the model can be solved. We specify that the household's period utility is given by

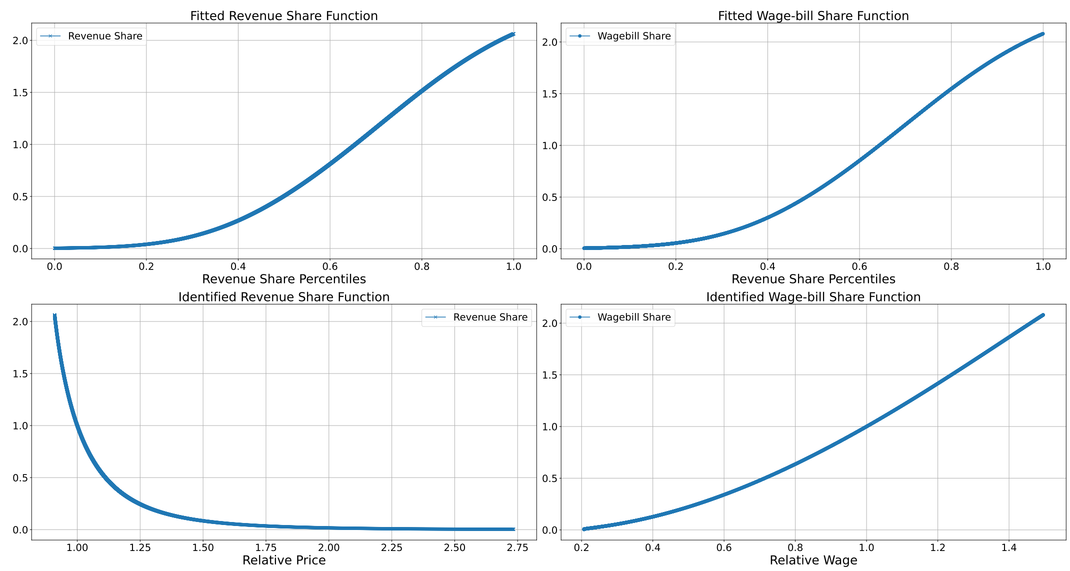
$$u(C_t, L_t) = \frac{C_t^{1-\gamma}-1}{1-\gamma} - \frac{L_t^{1+\varphi}}{1+\varphi}, \tag{41}$$

where $\gamma > 0$ is the elasticity of intertemporal substitution and $\varphi > 0$ is the inverse Frisch elasticity of labor supply. We use a standard calibration for these two parameters $\gamma = 1, \varphi = 1$.[40] Figure 9 shows the identified $r(\cdot)$ and $s(\cdot)$. The demand system $r(\cdot)$ seems to be more elastic than the labor supply $s(\cdot)$, consistent with our finding in the empirical section.

We assume two types of firms in the economy, distinguished by low and high quality-adjusted TFP. For each year, we calibrate two moments to characterize the firm distribution: the quality-adjusted TFP difference between low- and high-type firms, and the relative mass of high- and low-type firms. The first moment is calibrated using the empirically observed log difference in labor value-added TFP between the 90th and 10th percentiles. The second moment is determined by the relative mass of high- and low-type firms implied by the first moment and the observed TFP dispersion. Both the log difference between the 90th and 10th percentiles and the TFP dispersion are presented in Figure 7. Figure 10 shows the implied mass and TFP difference at the beginning and end of the sample. Note that, compared to 1977, the 2019 firm distribution features a greater share of high-productivity firms, and these firms are, on average, even more productive than their high-productivity counterparts in 1977.

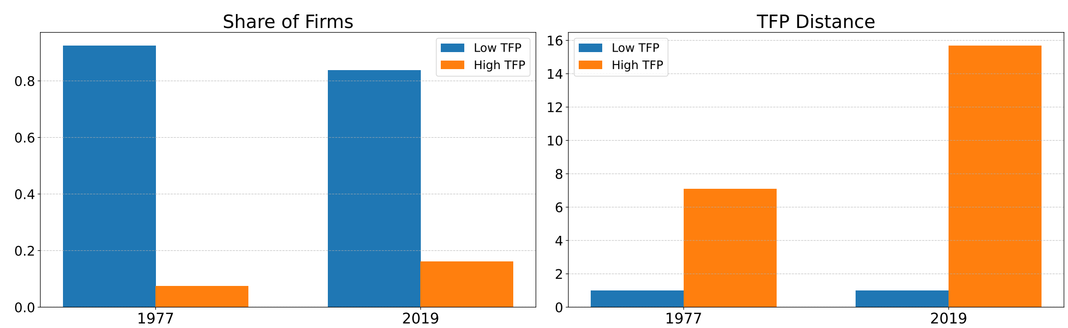
[40] Appendix F.2 provides more details on the solution method.

Figure 9: Identified Demand and Supply Systems



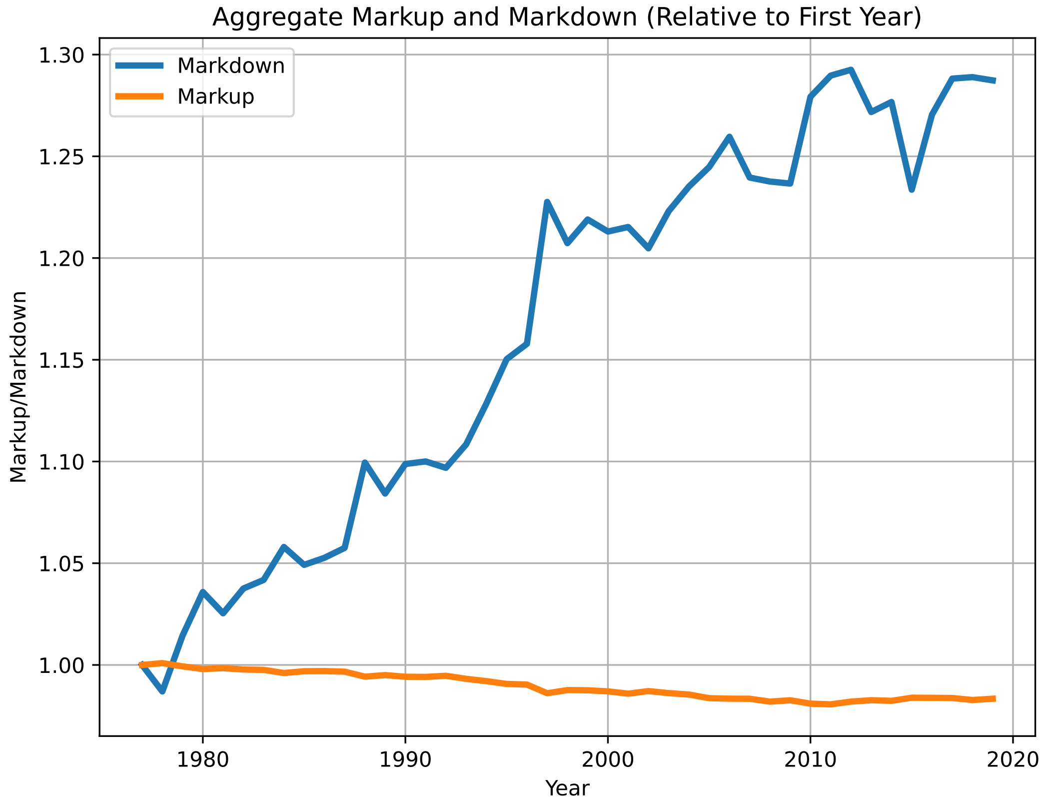
Notes: This figure presents the identified demand and labor supply systems using the joint distribution of markup, markdown, revenue share, and labor share observed in 1977. The first row displays the empirically fitted revenue share (left panel) and wage bill share (right panel) sorted by revenue share percentiles. The second panel row reports the identified revenue share (left panel) and wage bill share functions (right panel) as functions of relative prices and wages.

Figure 10: Firm Types Distribution



Notes: This figure presents the firm distribution for 1977 and 2019. The TFP distance is calibrated using the log difference in value-added labor productivity between the 90th and 10th percentiles. Based on this distance, the mass of different firm types is determined by matching the standard deviation of log labor productivity, as shown in Figure 7.

Figure 11: Evolution of the Model Implied Market Powers, 1977–2019



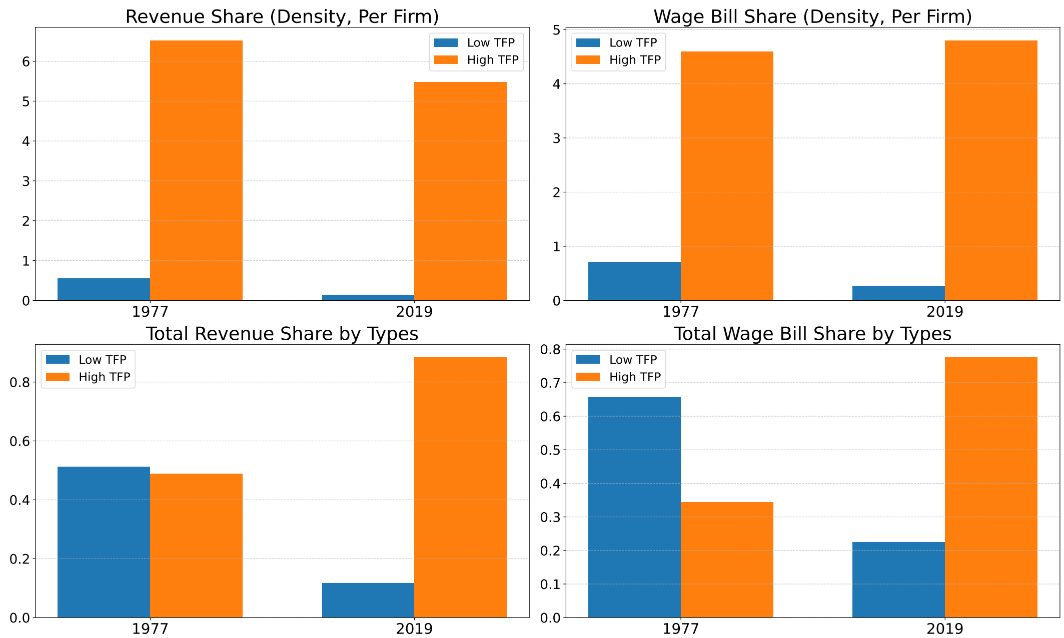
Notes: This figure reports the aggregate time series of markdowns and markups derived from the calibrated model. The model takes the demand and supply systems identified using 1977 data as given and solves for the steady state for each subsequent year, based on the firm distribution implied by the observed log 90th-to-10th percentile value-added labor productivity difference and the log value-added labor productivity dispersion. Aggregate markdowns are wage-bill-weighted, while aggregate markups are revenue-weighted. The series are normalized so that each point represents the cumulative growth relative to the first year.

6 Calibration Results

In this section, we begin by examining how changes in TFP dispersion drive increases (or decreases) in aggregate markups and markdowns. Next, we decompose these changes by analyzing the evolution of the cross-sectional distributions of revenue shares, wage bill shares, markups, and markdowns. Finally, we focus on pass-through, illustrating how our results align with existing findings and highlighting the implied changes in pass-through over time.

Figure 11 presents the model-implied time series of aggregate price markups and wage markdowns. Between 1977 and 2019, the aggregate markdown increased by approximately 30%, while the aggregate markup remained relatively stable. This pattern is qualitatively consistent with our empirical findings. Quantitatively, the increase in TFP dispersion accounts for roughly 30% of the

Figure 12: Decomposing Aggregate Trends: Market Share by Firm Types



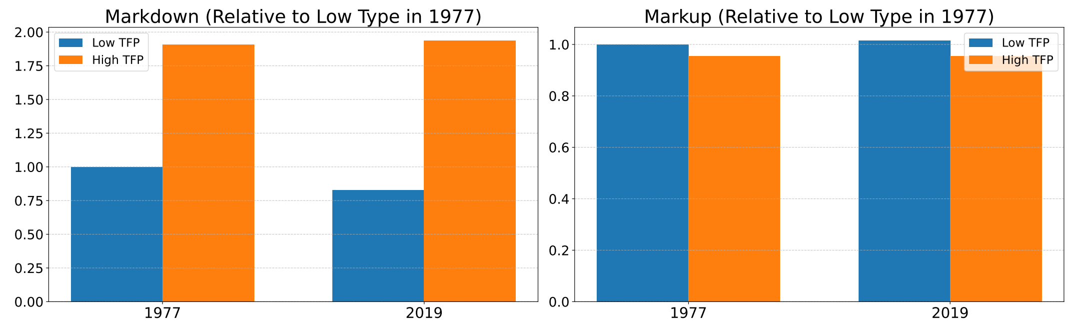
Notes: This figure reports the revenue and wage bill shares implied by the calibration. The first row shows the density of per-firm revenue shares (left panel) and wage bill shares (right panel), while the second row shows the total revenue (left panel) and wage bill shares (right panel) allocated to low- and high-TFP firms in 1977 and 2019. Blue bars represent the revenue and wage bill shares for low-TFP firms, while orange bars represent the corresponding shares for high-TFP firms.

observed rise in markdowns. This result is intuitive in light of the model's implications: as TFP dispersion increases, more productive firms expand their market reach, gain larger revenue shares, and reduce their markups. However, the resulting rise in product demand leads these firms to hire more labor, thereby enjoying greater markdowns. Our empirical cross-sectional results suggest that markups vary less across size or productivity than markdowns. That predicts a more stable time series evolution for markups relative to markdowns, which our model generates.

We now decompose the aggregate results by analyzing the cross-sectional distribution of revenue shares, wage bill shares, markups, and markdowns in 1977 and 2019. Figure 12 illustrates the revenue and wage bill shares of low- and high-productivity firms in these two years. Notably, the revenue shares of both low- and high-productivity firms individually decline over time. This pattern reflects the growing mass of high-productivity firms and the relatively elastic demand in product markets. The strong increase in competitive pressure decreases individual firm-level revenue shares. In contrast, the evolution of wage bill shares reveals a different dynamic. Since labor demand is less elastic, high-productivity firms are able to capture a greater wage bill share.

When examining total revenue shares across firm types, we observe a notable shift: compared to the relative parity in 1977, high-productivity firms now dominate, capturing the majority of the revenue share, underscoring the significant role of reallocation and matches the increase in

Figure 13: Decomposing Aggregate Trends: Market Power by Firm Types



Notes: This figure reports the wage markdown (left panel) and price markup (right panel) implied by the calibration. Both markups and markdowns are normalized to the values of low-TFP firms in 1977. Blue bars represent the markup and markdown for low-TFP firms, while orange bars represent the corresponding values for high-TFP firms.

concentration in the data. The changes in wage bill shares are even more striking. In 1977, low-productivity firms held a larger share of the total wage bill, but by 2019, high-productivity firms account for nearly 80% of the total wage bill share. This suggests that labor reallocation from low- to high-productivity firms has been even more pronounced than the reallocation of revenue shares.

Turning to firm-level markups and markdowns, Figure 13 displays the levels for low- and high-productivity firms in 1977 and 2019. For markdowns, low-productivity firms experience a slight decrease, while high-productivity firms see an increase. In contrast, markups increase for both low- and high-productivity firms. These results are consistent with the patterns observed in Figure 12: wage bill shares increase for high-productivity firms while decreasing for low-productivity firms, and revenue shares decline for both groups. However, the changes in both markup and markdown are relatively modest, especially compared to the aggregate increase in markdowns. This suggests that the bulk of the aggregate markdown increase is driven by reallocation effects.

Finally, we examine the model-implied pass-through. Figure 14 shows both primitive and effective pass-throughs. As implied in the data, primitive price pass-through exceeds 1, suggesting that larger firms face more elastic demand. Notably, low-productivity firms in 2019 exhibit lower pass-through than in 1977. Effective (measured) price pass-through, however, is incomplete. As explained by Proposition 4, this arises because firms operate under an upward-sloping labor supply curve. Two additional patterns emerge for the effective price pass-through: (1) high-productivity firms (larger firms) exhibit smaller effective price pass-through than low-productivity firms, and (2) the effective price pass-through of high-productivity firms is approximately half of that of low-productivity firms. This is consistent with the findings in Amiti, Itskhoki and Konings (2019). The lower effective price pass-through observed for high-productivity firms is explained by their exposure to a more inelastic labor supply curve, which amplifies the dampening effect of wage adjustments on price changes. Finally, both primitive and effective wage pass-through are incomplete, consistent with prior empirical findings.

Figure 14: Pass-through by Firm Types

Notes: This figure reports the primitive and effective pass-throughs implied by the calibration. The first row illustrates the primitive price (left panel) and wage pass-throughs (right panel), while the second row highlights the effective price (left panel) and wage pass-throughs (right panel). Blue bars represent the pass-through measures for low-TFP firms, while orange bars represent the corresponding measures for high-TFP firms.

7 Conclusion

Our study provides a new perspective on the rise of aggregate market power in the United States by disentangling product and labor market power. While previous research has primarily focused on increasing price markups, our analysis reveals that the secular increase in aggregate market power has been largely driven by growing wage markdowns and labor market power, rather than rising price markups. Empirically, we document that the aggregate price markup has remained stable, while the aggregate wage markdown has increased significantly, nearly doubling over the last few decades. This trend aligns with the broader economic shift toward increasing productivity dispersion and the rising dominance of high-productivity firms, suggesting that reallocation effects are a key driver of observed changes in market power.

Furthermore, our decomposition of market power provides a novel perspective on the relationship between firm size and market power. Contrary to the standard assumption of Marshall's Second Law of Demand, which implies that larger firms exhibit higher markups and therefore exhibit incomplete price pass-throughs, we find that markups are negatively correlated with firm size. Conversely, markdowns exhibit a strong positive relationship. We also show that labor market power, rather than product market power, is the primary factor behind the positive correlation between firm size and total market power. Our structural model reconciles our empirical results with prior empirical findings of measured incomplete price pass-through. It shows that the interaction between labor market power and product market power can generate measured incomplete pass-throughs even when Marshall's Second Law of Demand does not hold.

Our findings have broad implications for both policy and economic theory. From a policy perspective, the increasing concentration of labor market power raises important concerns about wage stagnation and labor market inequality. Standard antitrust policies, which primarily target product market power, may need to be reconsidered to account for monopsony power in labor markets. Furthermore, our evidence suggests that the efficiency gains often associated with firm expansion may come at the cost of increasing wage suppression, highlighting the need for a more nuanced approach to regulating market power.

Theoretically, our results challenge the assumption that incomplete price pass-through is primarily a function of product market competition and how markups vary across size. Instead, we show that firms' market power in input markets plays a crucial role in shaping pricing behavior. This insight has important ramifications for macroeconomic models that rely on firm-level price-setting behavior, as it suggests that labor market structures should be incorporated into standard frameworks to fully capture the dynamics of market power.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica*, 83(6): 2411–2451.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li. 2023. "A Theory of Falling Growth and Rising Rents." *Review of Economic Studies*, 90(6): 2675–2702.

Akcigit, Ufuk, and Sina T. Ates. 2021. "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory." *American Economic Journal: Macroeconomics*, 13(1): 257–298.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2019. "International Shocks, Variable Markups, and Domestic Prices." *Review of Economic Studies*, 86(6): 2356–2402.

Anderson, Eric, Sergio Rebelo, and Arlene Wong. 2018. "Markups Across Space and Time." NBER Working Paper 24434.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal. 2015. "Frontier Firms, Technology Diffusion and Public Policy: Micro Evidence from OECD Countries." OECD Productivity Working Papers.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal. 2016. "The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy." OECD Productivity Working Papers.

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review*, 98(5): 1998–2031.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. "Concentrating on the Fall of the Labor Share." *American Economic Review Papers & Proceedings*, 107(5): 180–185.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *Quarterly Journal of Economics*, 135(2): 645–709.

Baqaee, David R., Emmanuel Farhi, and Kunal Sangani. 2024*a*. "The Supply-Side Effects of Monetary Policy." *Journal of Political Economy*, 132(4): 1065–1112.

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *Quarterly Journal of Economics*, 135(1): 105–163.

Baqaee, David Rezza, Emmanuel Farhi, and Kunal Sangani. 2024*b*. "The Darwinian Returns to Scale." *Review of Economic Studies*, 91(3): 1373–1405.

Barkai, Simcha. 2020. "Declining Labor and Capital Shares." *Journal of Finance*, 75(5): 2421–2463.

Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman. 2016. "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States." *Journal of Labor Economics*, 34(S2): S67–S97.

Basu, Susanto, and John G. Fernald. 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy*, 105(2): 249–283.

Benmelech, Efraim, Nittai K. Bergman, and Hyunseob Kim. 2022. "Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?" *Journal of Human Resources*, 57(S): S200–S250.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor Market Power." *American Economic Review*, 112(4): 1147–1193.

Berger, David W., Kyle F. Herkenhoff, Andreas R. Kostøl, and Simon Mongey. 2023. "An Anatomy of Monopsony: Search Frictions, Amenities, and Bargaining in Concentrated Markets." In *NBER Macroeconomics Annual*. Vol. 38. University of Chicago Press.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica*, 77(3): 623–685.

Bloom, Nicholas, Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter. 2018. "The Disappearing Large-Firm Wage Premium." *AEA Papers and Proceedings*, 108: 317–322.

Board of Governors of the Federal Reserve System. 2024. "H.15 Selected Interest Rates (Daily)."

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation From Production Data." *Journal of Monetary Economics*, 121: 1–14.

Brooks, Wyatt J., Joseph P. Kaboski, Yao Amber Li, and Wei Qian. 2021. "Exploitation of Labor? Classical Monopsony Power and Labor's Share." *Journal of Development Economics*, 150: 102627.

Brown, Charles, and James Medoff. 1989. "The Employer Size-Wage Effect." *Journal of Political Economy*, 97(5): 1027–1059.

Bureau of Economic Analysis. 2024. "National Income and Product Accounts."

Bureau of Labor Statistics. 2024. "Consumer Price Index."

Burstein, Ariel, and Gita Gopinath. 2014. "Chapter 7 - International Prices and Exchange Rates." In *Handbook of International Economics*. Vol. 4, , ed. Gita Gopinath, Elhanan Helpman and Kenneth Rogoff, 391–451. Elsevier.

Bustamante, M. Cecilia, and Andres Donangelo. 2017. "Product Market Competition and Industry Returns." *Review of Financial Studies*, 30(12): 4216–4266.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S70.

Cavenaile, Laurent, and Pau Roldan-Blanco. 2021. "Advertising, Innovation, and Economic Growth." *American Economic Journal: Macroeconomics*, 13(3): 251–303.

Center for Research in Security Prices. 2020. "CRSP/Compustat Merged Database."

Chan, Mons, Sergio Salgado, and Ming Xu. 2023. "Heterogeneous Passthrough from TFP to Wages." SSRN Working Paper.

Chow, Melissa, Teresa C. Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White. 2021. "Redesigning the Longitudinal Business Database." U.S. Census Bureau CES 21-08.

Cooper, Russell, John Haltiwanger, and Jonathan L. Willis. 2007. "Search Frictions: Matching Aggregate and Establishment Observations." *Journal of Monetary Economics*, 54: 56–78.

Cooper, Russell W., and John C. Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies*, 73(3): 611–633.

Corhay, Alexandre, Howard Kung, and Lukas Schmid. 2020. "Competition, Markups, and Predictable Returns." *Review of Financial Studies*, 33(12): 5906–5939.

Crouzet, Nicolas, and Janice Eberly. 2021. "Intangibles, Markups, and the Measurement of Productivity Growth." *Journal of Monetary Economics*, 124: S92–S109.

Crouzet, Nicolas, and Janice Eberly. 2023. "Rents and Intangible Capital: A Q+ Framework." *Journal of Finance*, 78(4): 1873–1916.

David, Joel M., and Venky Venkateswaran. 2019. "The Sources of Capital Misallocation." *American Economic Review*, 109(7): 2531–2567.

Delabastita, Vincent, and Michael Rubens. 2024. "Colluding Against Workers." SSRN Working Paper.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review*, 102(6): 2437–2471.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics*, 135(2): 561–644.

Demirer, Mert. 2022. "Production Function Estimation with Factor-Augmenting Technology: An Application to Markups." Working Paper.

De Ridder, Maarten. 2024. "Market Power and Innovation in the Intangible Economy." *American Economic Review*, 114(1): 199–251.

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti. 2025. "The Hitchhiker's Guide to Markup Estimation." Working Paper.

Díez, Federico J., Jiayue Fan, and Carolina Villegas-Sánchez. 2021. "Global Declining Competition?" *Journal of International Economics*, 132: 103492.

Dobbelaere, Sabien, and Jacques Mairesse. 2013. "Panel Data Estimates of the Production Function and Product and Labor Market Imperfections." *Journal of Applied Econometrics*, 28(1): 1–46.

Donangelo, Andres, François Gourio, Matthias Kehrig, and Miguel Palacios. 2019. "The Cross-Section of Labor Leverage and Equity Returns." *Journal of Financial Economics*, 132(2): 497–518.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2019. "Using Cost Minimization to Estimate Markups." Working Paper.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015. "Competition, Markups, and the Gains from International Trade." *American Economic Review*, 105(10): 3183–3221.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2023. "How Costly Are Markups?" *Journal of Political Economy*, 131(7): 1619–1675.

Eisfeldt, Andrea L., and Dimitris Papanikolaou. 2013. "Organization Capital and the Cross-Section of Expected Returns." *Journal of Finance*, 68(4): 1365–1406.

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin. 2013. "The Decline of the U.S. Labor Share." *Brookings Papers on Economic Activity*, 44(2): 1–52.

Feenstra, Robert C., Joseph E. Gagnon, and Michael M. Knetter. 1996. "Market Share and Exchange Rate Pass-Through in World Automobile Trade." *Journal of International Economics*, 40(1): 187–207.

Flynn, Zach, James Traina, and Amit Gandhi. 2019. "Measuring Markups with Production Data." SSRN Working Paper.

Gandhi, Amit, Salvador Navarro, and David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." *Journal of Political Economy*, 128(8): 2973–3016.

Goldberg, Pinelopi Koujianou, and Rebecca Hellerstein. 2013. "A Structural Approach to Identifying the Sources of Local Currency Price Stability." *Review of Economic Studies*, 80(1): 175–210.

Gouin-Bonenfant, Émilien. 2022. "Productivity Dispersion, Between-Firm Competition, and the Labor Share." *Econometrica*, 90(6): 2755–2793.

Grossman, Gene M., Elhanan Helpman, and Hugo Lhuillier. 2023. "Supply Chain Resilience: Should Policy Promote International Diversification or Reshoring?" *Journal of Political Economy*, 131(12): 3462–3496.

Gutiérrez, Germán, and Thomas Philippon. 2017. "Declining Competition and Investment in the U.S." NBER Working Paper 23583.

Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy*, 96(5): 921–947.

Hall, Robert E. 2004. "Measuring Factor Adjustment Costs." *Quarterly Journal of Economics*, 119(3): 899–927.

Hurst, Erik, Patrick J. Kehoe, Elena Pastorino, and Thomas Winberry. 2022. "The Distributional Impact of the Minimum Wage in the Short and Long Run." National Bureau of Economic Research NBER Working Paper 30294.

Karabarbounis, Loukas, and Brent Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics*, 129(1): 61–103.

Kehrig, Matthias. 2015. "The Cyclical Nature of the Productivity Distribution." SSRN Working Paper.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar. 2019. "Who Profits from Patents? Rent-Sharing at Innovative Firms." *Quarterly Journal of Economics*, 134(3): 1343–1404.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler. 2023. "Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry." NBER Working Paper 27325.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2022. "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market." *American Economic Review*, 112(1): 169–212.

Lehr, Nils H. 2023. "Does Monopsony Matter for Innovation?" Working Paper.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*, 70(2): 317–341.

Lipsius, Ben. 2018. "Labor Market Concentration Does Not Explain the Falling Labor Share." SSRN Working Paper.

Mankiw, N. Gregory, and Michael D. Whinston. 1986. "Free Entry and Social Inefficiency." *RAND Journal of Economics*, 17(1): 48–58.

Matsuyama, Kiminori. 2023. “Non-CES Aggregators: A Guided Tour.” *Annual Review of Economics*, 15: 235–265.

Matsuyama, Kiminori. 2025. “Homothetic Non-CES Demand Systems with Applications to Monopolistic Competition.” *Annual Review of Economics*, 17.

Matsuyama, Kiminori, and Philip Ushchev. 2017. “Beyond CES: Three Alternative Classes of Flexible Homothetic Demand Systems.” Working Paper.

Matsuyama, Kiminori, and Philip Ushchev. 2021*a*. “Constant Pass-Through.” Working Paper.

Matsuyama, Kiminori, and Philip Ushchev. 2021*b*. “When Does Procompetitive Entry Imply Excessive Entry.” Working Paper.

Matsuyama, Kiminori, and Philip Ushchev. 2022. “Destabilizing Effects of Market Size in the Dynamics of Innovation.” *Journal of Economic Theory*, 200: 105415.

Matsuyama, Kiminori, and Philip Ushchev. 2023*a*. “Love-For-Variety.” Working Paper.

Matsuyama, Kiminori, and Philip Ushchev. 2023*b*. “Selection and Sorting of Heterogeneous Firms through Competitive Pressures.” Working Paper.

Melitz, Marc. 2018. “Trade Competition and Reallocations in a Small Open Economy.” *World Trade Evolution*, 60–81.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71(6): 1695–1725.

Mertens, Matthias, and Bernardo Mottironi. 2023. “Do Larger Firms Exert More Market Power? Markups and Markdowns Along the Size Distribution.” Working Paper.

Morlacco, Monica. 2020. “Market Power in Input Markets: Theory and Evidence from French Manufacturing.” Working Paper.

Mrázová, Monika, and J. Peter Neary. 2017. “Not So Demanding: Demand Structure and Firm Behavior.” *American Economic Review*, 107(12): 3835–3874.

Mrázová, Monika, and J. Peter Neary. 2020. “IO for Exports(s).” *International Journal of Industrial Organization*, 70: 102561.

Nakamura, Emi, and Dawit Zerom. 2010. “Accounting for Incomplete Pass-Through.” *Review of Economic Studies*, 77(3): 1192–1230.

Olley, G. Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, 64(6): 1263–1297.

Peters, Michael. 2020. "Heterogeneous Markups, Growth, and Endogenous Misallocation." *Econometrica*, 88(5): 2037–2073.

Peters, Ryan H., and Lucian A. Taylor. 2017. "Intangible Capital and the Investment-q Relation." *Journal of Financial Economics*, 123(2): 251–272.

Raval, Devesh. 2023*a*. "A Flexible Cost Share Approach to Markup Estimation." *Economics Letters*, 230: 111262.

Raval, Devesh. 2023*b*. "Testing the Production Approach to Markup Estimation." *Review of Economic Studies*, 90(5): 2592–2611.

Rubens, Michael. 2023. "Market Structure, Oligopsony Power, and Productivity." *American Economic Review*, 113(9): 2382–2410.

Sangani, Kunal. 2023. "Pass-Through in Levels and the Incidence of Commodity Shocks." Working Paper.

Seegmiller, Bryan. 2023. "Valuing Labor Market Power: The Role of Productivity Advantages." Working Paper.

Staiger, Douglas O., Joanne Spetz, and Ciaran S. Phibbs. 2010. "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment." *Journal of Labor Economics*, 28(2): 211–236.

Syverson, Chad. 2004*a*. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy*, 112(6): 1181–1222.

Syverson, Chad. 2004*b*. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics*, 86(2): 534–550.

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature*, 49(2): 326–365.

Traina, James. 2018. "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements." SSRN Working Paper.

Trottner, Fabian. 2023. "Unbundling Market Power." Working Paper.

United States Census Bureau. 2022. "Longitudinal Business Database."

Wang, Olivier, and Iván Werning. 2022. "Dynamic Oligopoly and Price Stickiness." *American Economic Review*, 112(8): 2815–2849.

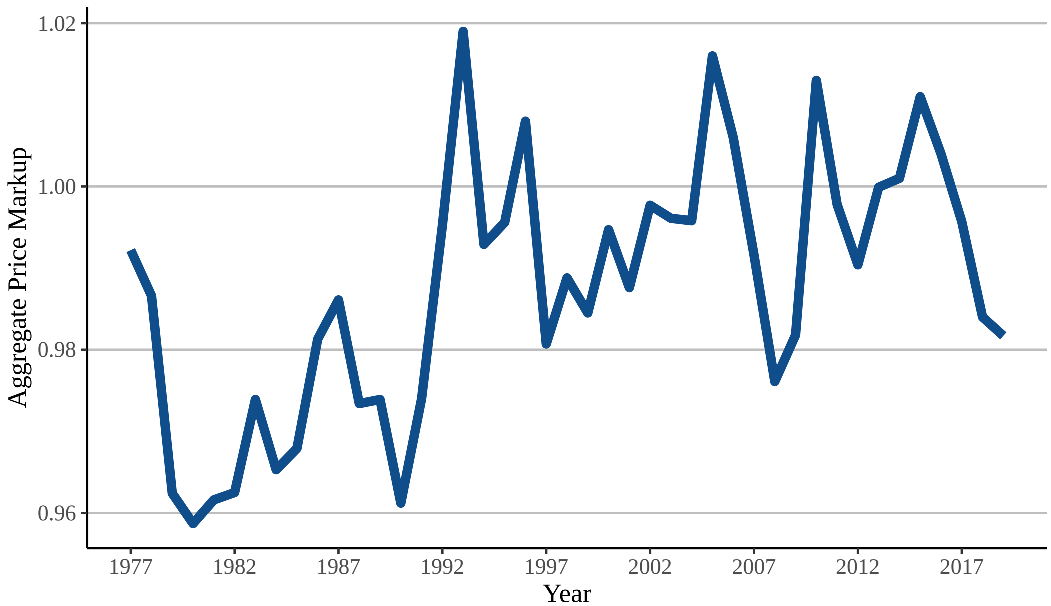
Webber, Douglas A. 2015. "Firm Market Power and the Earnings Distribution." *Labour Economics*, 35: 123–134.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein. 2022. "Monopsony in the US Labor Market." *American Economic Review*, 112(7): 2099–2138.

A Additional Figures and Tables

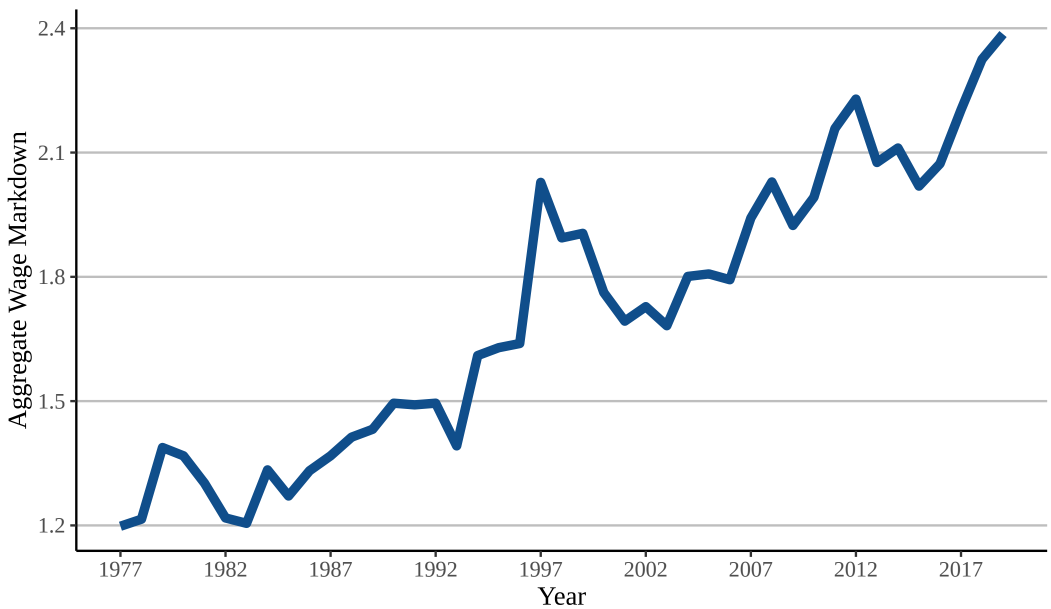
A.1 Figures

Figure A1: Evolution of the Aggregate Price Markup (Levels), 1977–2019



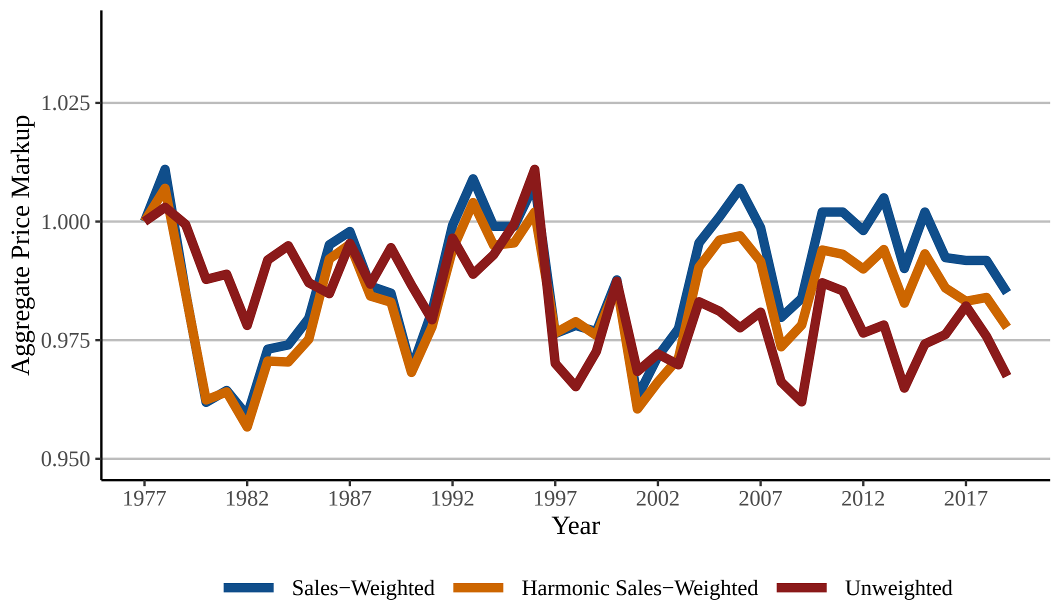
Notes: This figure reports the aggregate price markup following the definition in Equation (11) from 1977 to 2019 in levels. Production functions are estimated using a CRS restriction, and the aggregate price markup is computed using sales as the weight. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A2: Evolution of the Aggregate Wage Markdown (Levels), 1977–2019



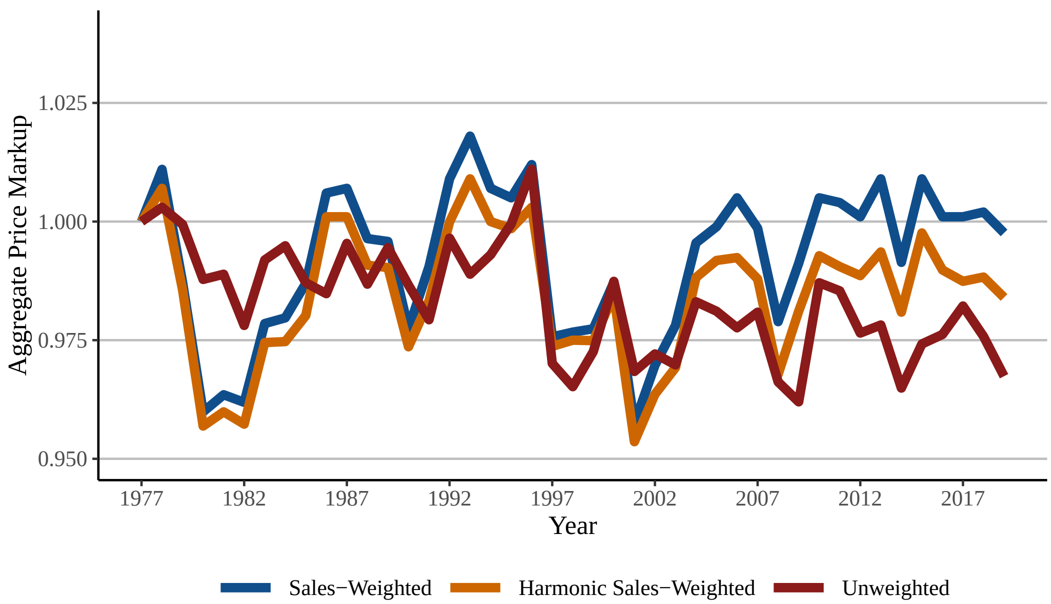
Notes: This figure reports the aggregate wage markdown following the definition in Equation (12) from 1977 to 2019 in levels. Production functions are estimated using a CRS restriction, and the aggregate wage markdown is computed using the wage bill as the weight. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A3: Aggregate Price Markup Indices under Various Weighting Methodologies (CRS Restriction)



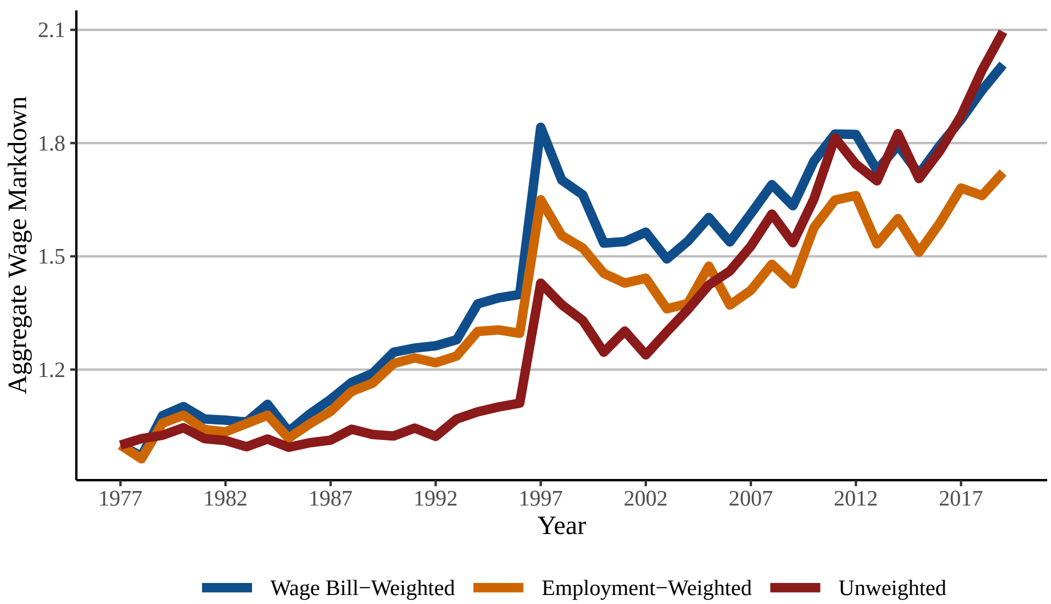
Notes: This figure displays index versions of the aggregate price markup under different weighting methodologies: sales-weighted (blue), harmonic average (orange), and simple average/unweighted (red). The series are normalized to 1 in 1977, and the estimates cover 1977 to 2019. In these specifications, all production functions are estimated using the CRS restriction. These series are the index counterpart to those in Figure A7. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A4: Aggregate Price Markup Indices under Various Weighting Methodologies (No CRS Restriction)



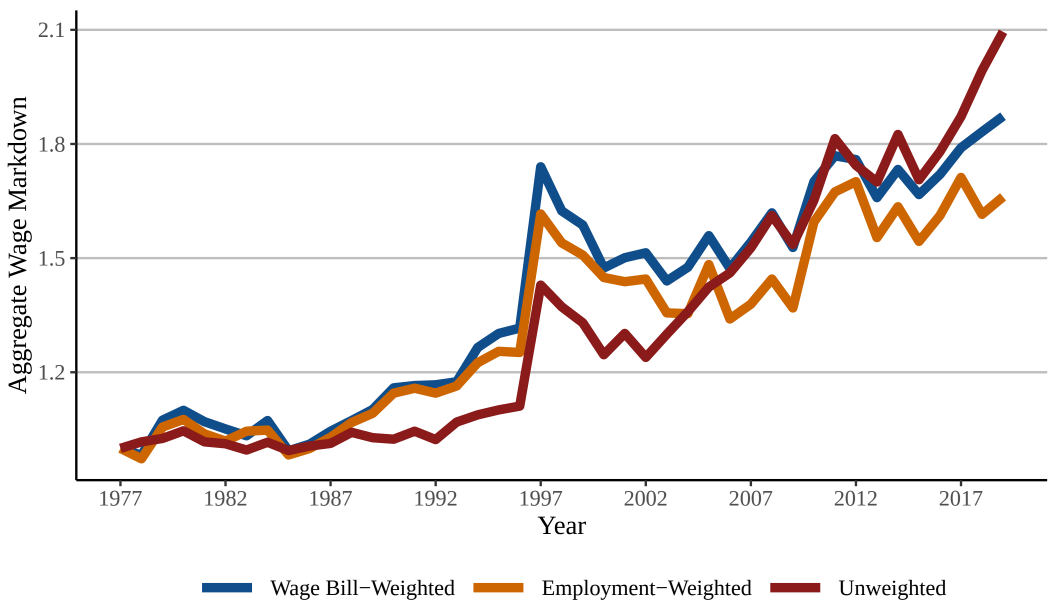
Notes: This figure displays index versions of the aggregate price markup under different weighting methodologies: sales-weighted (blue), harmonic average (orange), and simple average/unweighted (red). The series are normalized to 1 in 1977, and the estimates cover 1977 to 2019. In these specifications, all production functions are estimated without using the CRS restriction. These series are the index counterpart to those in Figure A8. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A5: Aggregate Wage Markdown Indices under Various Weighting Methodologies (CRS Restriction)



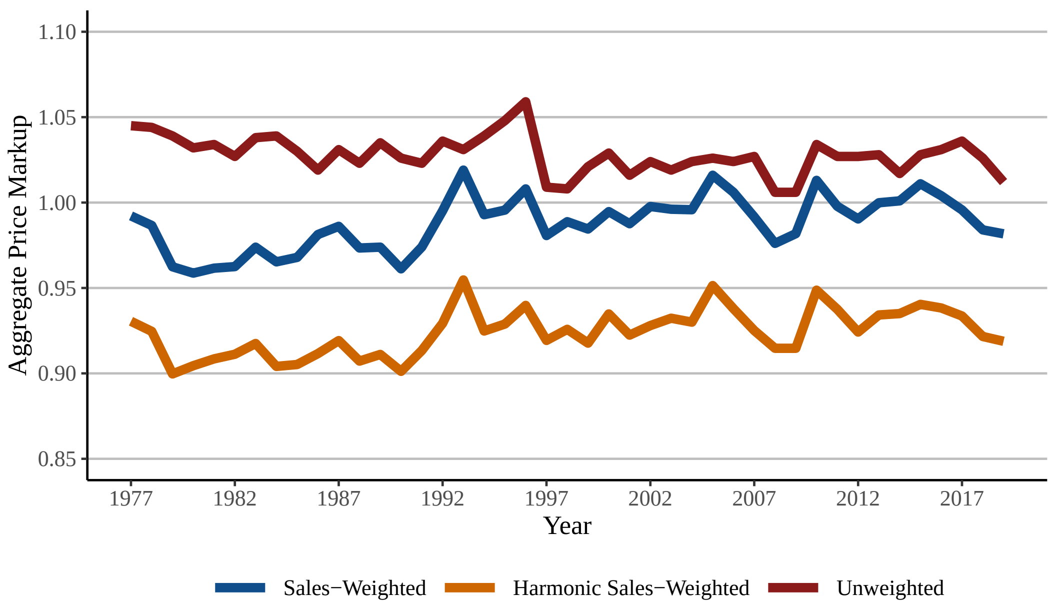
Notes: This figure displays index versions of the aggregate wage markdown under different weighting methodologies: wage bill-weighted (blue), employment-weighted (orange), and simple average/unweighted (red). The series are normalized to 1 in 1977, and the estimates cover 1977 to 2019. In these specifications, all production functions are estimated using the CRS restriction. These series are the index counterpart to those in Figure A9. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A6: Aggregate Wage Markdown Indices under Various Weighting Methodologies (No CRS Restriction)



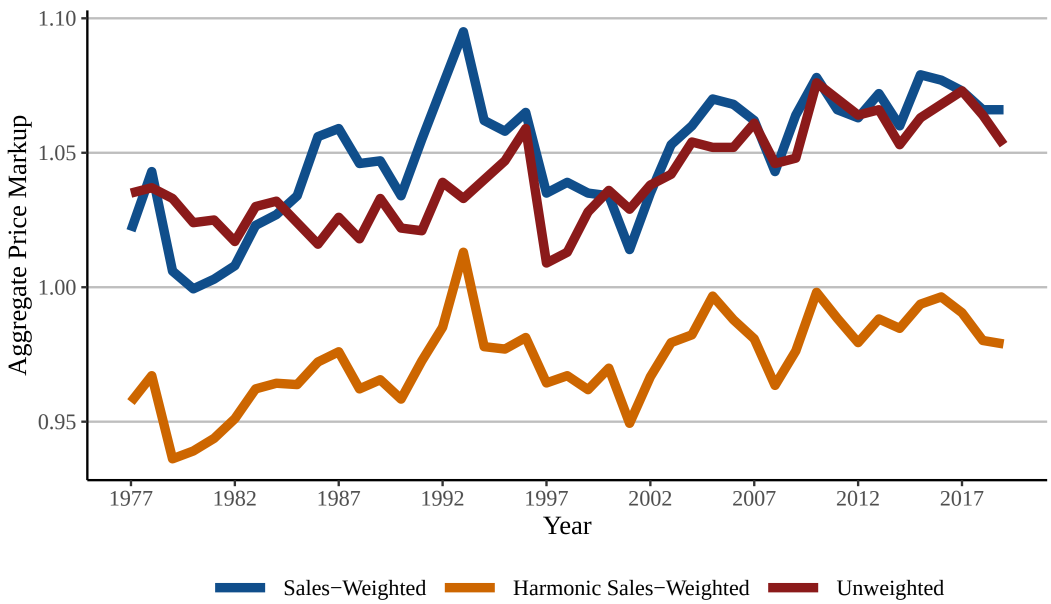
Notes: This figure displays index versions of the aggregate wage markdown under different weighting methodologies: wage bill-weighted (blue), employment-weighted (orange), and simple average/unweighted (red). The series are normalized to 1 in 1977, and the estimates cover 1977 to 2019. In these specifications, all production functions are estimated without using the CRS restriction. These series are the index counterpart to those in Figure A10. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A7: Aggregate Price Markup under Various Weighting Methodologies (CRS Restriction)



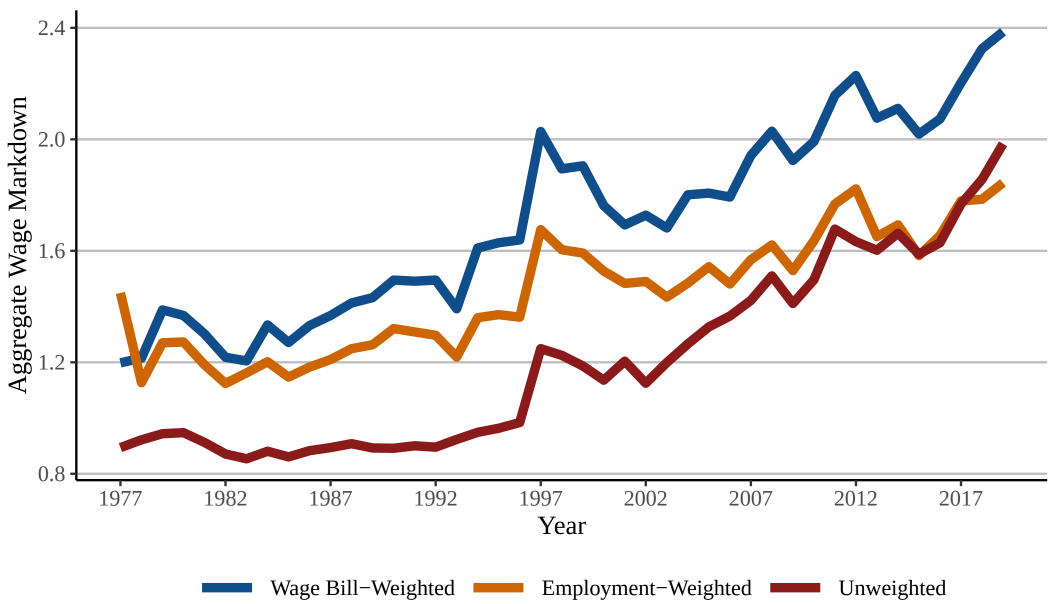
Notes: This figure displays the aggregate price markup under different weighting methodologies: sales-weighted (blue), harmonic average (orange), and simple average/unweighted (red). The estimates cover 1977 to 2019. In these specifications, all production functions are estimated using the CRS restriction. These series are the levels counterpart to those in Figure A3. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A8: Aggregate Price Markup under Various Weighting Methodologies (No CRS Restriction)



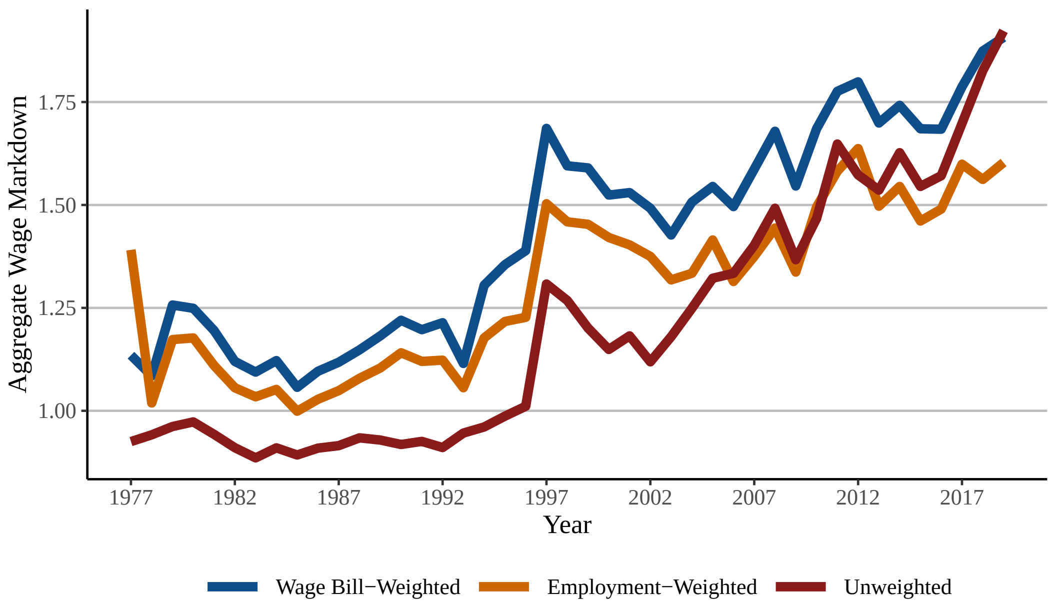
Notes: This figure displays the aggregate price markup under different weighting methodologies: sales-weighted (blue), harmonic average (orange), and simple average/unweighted (red). The estimates cover 1977 to 2019. In these specifications, all production functions are estimated without using the CRS restriction. These series are the levels counterpart to those in Figure A4. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A9: Aggregate Wage Markdown under Various Weighting Methodologies (CRS Restriction)



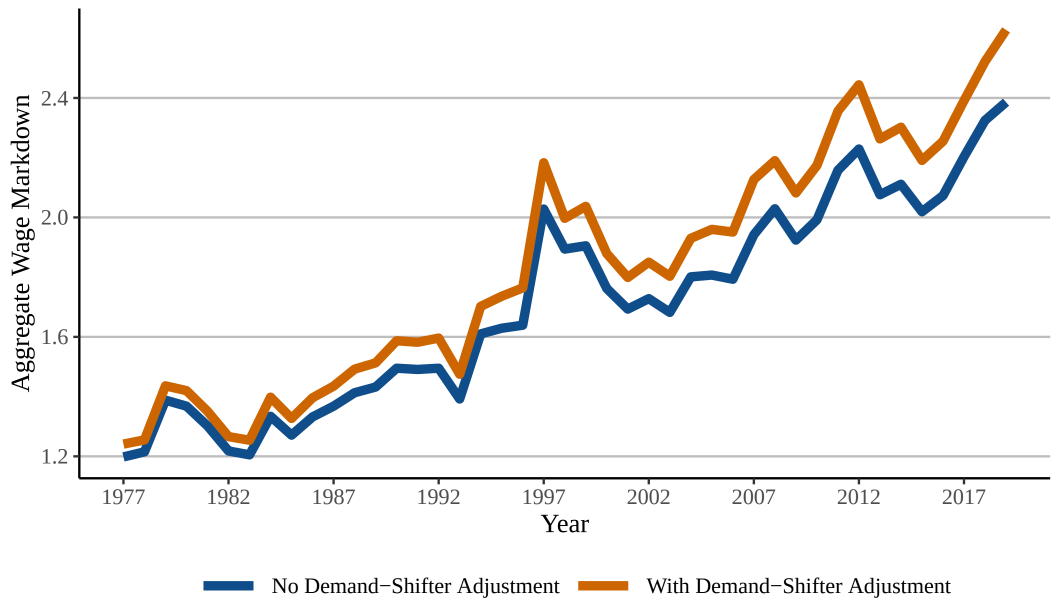
Notes: This figure displays the aggregate wage markdown under different weighting methodologies: wage bill-weighted (blue), employment-weighted (orange), and simple average/unweighted (red). The estimates cover 1977 to 2019. In these specifications, all production functions are estimated using the CRS restriction. These series are the levels counterpart to those in Figure A5. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A10: Aggregate Wage Markdown under Various Weighting Methodologies (No CRS Restriction)



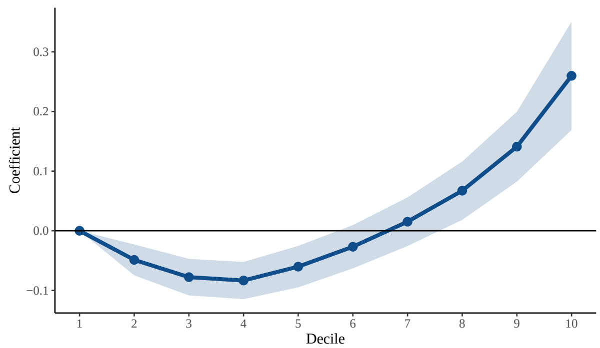
Notes: This figure displays the aggregate wage markdown under different weighting methodologies: wage bill-weighted (blue), employment-weighted (orange), and simple average/unweighted (red). The estimates cover 1977 to 2019. In these specifications, all production functions are estimated without using the CRS restriction. These series are the levels counterpart to those in Figure A6. All figures are rounded in accordance with U.S. Census disclosure requirements.

Figure A11: Aggregate Wage Markdown with and without a Demand-Shifter Adjustment (CRS Restriction)

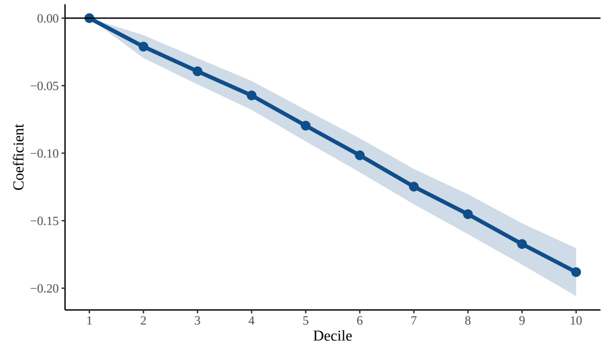


Notes: This figure displays the aggregate wage markdown in a baseline specification (blue) and a demand-shifter-adjusted specification (orange). The estimates cover 1977 to 2019. In these specifications, all production functions are estimated using the CRS restriction. See Appendix C.3 for details. All figures are rounded in accordance with U.S. Census disclosure requirements.

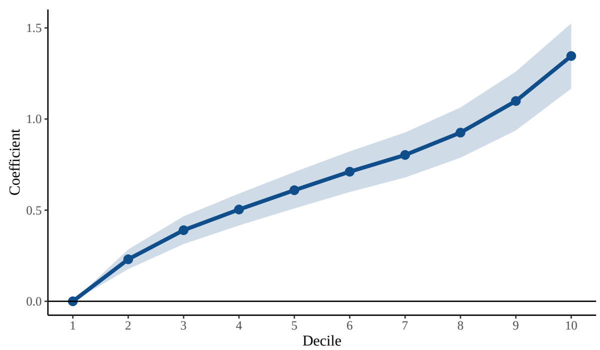
Figure A12: Markup and Markdown Across Deciles of Firm Size Measures (with Firm Fixed Effects)



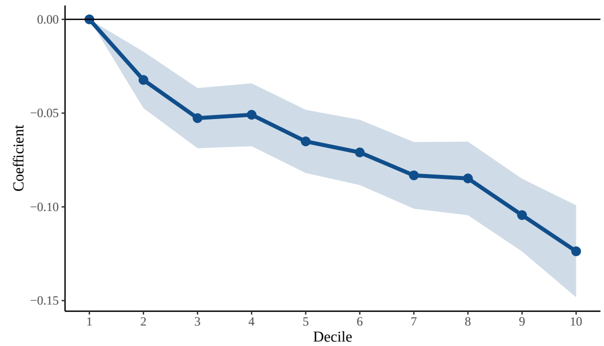
(a) Markdown Across Deciles of Log Labor Prod



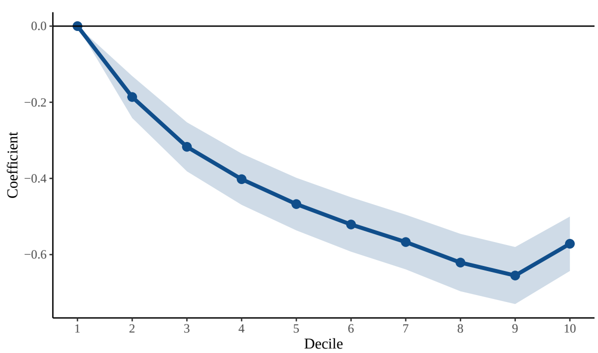
(b) Markup Across Deciles of Log Labor Prod



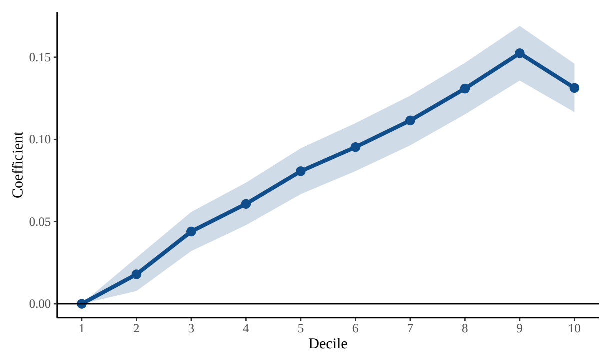
(c) Markdown Across Deciles of Log Sales



(d) Markup Across Deciles of Log Sales



(e) Markdown Across Deciles of Log Labor Share



(f) Markup Across Deciles of Log Labor Share

Notes: This figure reports the relationship between price markups and wage markdowns and various measures of firm size, as well as value-added labor share. Each panel reports the coefficients of the decile dummies from estimating Equation (16) with NAICS2 $\times$ year and firm fixed effects. Panels (a) and (b) present the coefficients for log value-added labor productivity. Panels (c) and (d) display the coefficients for log revenue. Panels (e) and (f) display the coefficients for log value-added labor shares. All figures are rounded in accordance with U.S. Census disclosure requirements.

A.2 Tables

Table A1: NAICS2-Level Aggregate Price Markups and Wage Markdowns Growth

NAICS2 Industry (1)	NAICS2 Code (2)	Markup Growth (%) (3)	Markdown Growth (%) (4)
Agriculture	11	2.3	58.4
Mining	21	26.0	11.0
Construction	23	1.8	2.3
Manufacturing I	31	2.3	101.2
Manufacturing II	32	6.8	170.2
Manufacturing III	33	-3.4	101.7
Wholesale Trade	42	-3.4	115.4
Retail Trade I	44	2.3	28.3
Retail Trade II	45	-10.6	99.6
Transport/Warehouse I	48	-15.0	114.6
Transport/Warehouse II	49	-6.2	124.8
Information	51	8.5	48.1
Professional Services	54	-4.5	0.8
Admin./Waste Services	56	-6.6	25.1
Education	61	18.3	-26.2
Healthcare/Social	62	13.5	19.2
Arts/Entertainment	71	-12.7	58.8
Accommodation/Food	72	10.2	-40.1
Other Services	81	8.1	-1.8

Notes: This table shows the NAICS2-level growth in the aggregate price markup and wage markdown from the first to the last decade of the sample. Column (1) contains the industry name and Column (2) contains the NAICS2 code. Columns (3) and (4) show the sales-weighted NAICS2-level aggregate price markup growth and the wage bill-weighted NAICS2-level aggregate wage markdown growth, respectively. The production functions are estimated using the CRS restriction. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A2: Price Markups and Wage Markdowns Summary Statistics (1977)

Variable	Mean (1)	SD (2)	P10 (3)	P25 (4)	Median (5)	P75 (6)	P90 (7)
Price Markup (CRS)	1.045	0.274	0.712	0.798	1.042	1.223	1.386
Wage Markdown (CRS)	0.894	0.838	0.199	0.338	0.584	1.167	2.007
Wage Markdown (CRS, DS Adj.)	0.938	0.892	0.203	0.351	0.626	1.206	2.041
Price Markup (No CRS)	1.035	0.250	0.717	0.806	1.019	1.207	1.367
Wage Markdown (No CRS)	0.925	0.718	0.289	0.438	0.717	1.177	1.845
Wage Markdown (No CRS, DS Adj.)	0.967	0.752	0.299	0.450	0.768	1.233	1.891

Notes: This table presents the summary statistics of the estimated price markups and wage markdowns in 1977 only. The table includes the estimates with and without a CRS restriction. The wage markdown estimates also include a version with a demand-shifter adjustment. Column (1) reports the mean, and Column (2) reports the standard deviation. Columns (3) to (7) report the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile, respectively. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A3: Price Markups and Wage Markdowns Summary Statistics (2000)

Variable	Mean (1)	SD (2)	P10 (3)	P25 (4)	Median (5)	P75 (6)	P90 (7)
Price Markup (CRS)	1.029	0.322	0.683	0.747	1.023	1.195	1.482
Wage Markdown (CRS)	1.136	1.259	0.223	0.381	0.684	1.395	2.599
Wage Markdown (CRS, DS Adj.)	1.184	1.339	0.229	0.393	0.702	1.443	2.699
Price Markup (No CRS)	1.036	0.293	0.697	0.769	1.000	1.207	1.467
Wage Markdown (No CRS)	1.149	1.137	0.273	0.446	0.762	1.458	2.448
Wage Markdown (No CRS, DS Adj.)	1.196	1.209	0.279	0.461	0.791	1.492	2.479

Notes: This table presents the summary statistics of the estimated price markups and wage markdowns in 2000 only. The table includes the estimates with and without a CRS restriction. The wage markdown estimates also include a version with a demand-shifter adjustment. Column (1) reports the mean, and Column (2) reports the standard deviation. Columns (3) to (7) report the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile, respectively. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A4: Price Markups and Wage Markdowns Summary Statistics (2019)

Variable	Mean (1)	SD (2)	P10 (3)	P25 (4)	Median (5)	P75 (6)	P90 (7)
Price Markup (CRS)	1.012	0.328	0.643	0.705	1.015	1.185	1.430
Wage Markdown (CRS)	1.984	2.378	0.296	0.552	1.056	2.466	4.729
Wage Markdown (CRS, DS Adj.)	2.104	2.517	0.310	0.601	1.147	2.682	4.959
Price Markup (No CRS)	1.053	0.342	0.654	0.726	1.026	1.241	1.559
Wage Markdown (No CRS)	1.923	2.220	0.307	0.547	1.096	2.494	4.466
Wage Markdown (No CRS, DS Adj.)	2.026	2.326	0.319	0.586	1.186	2.637	4.669

Notes: This table presents the summary statistics of the estimated price markups and wage markdowns in 2019 only. The table includes the estimates with and without a CRS restriction. The wage markdown estimates also include a version with a demand-shifter adjustment. Column (1) reports the mean, and Column (2) reports the standard deviation. Columns (3) to (7) report the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile, respectively. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A5: Comparison of Markup Estimates

	Log Markup (CRS)		
	(1)	(2)	(3)
Intercept	-0.015		
	(0.003)		
Log Markup (Without CRS)	0.933	0.934	0.948
	(0.011)	(0.004)	(0.003)
NAICS2 × Year FE	No	Yes	Yes
Firm FE	No	No	Yes
Observations	69,500	69,500	69,500

Notes: This table reports the regression results from regressing the markup estimate with the CRS restriction onto the markup estimate without the CRS restriction. Columns (1) to (3) report the result without any fixed effects, with NAICS2 × year fixed effects, and with both NAICS2 × year and firm fixed effects. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A6: Comparison of Markdown Estimates

	Log Markdown (CRS)		
	(1)	(2)	(3)
Intercept	-0.102		
	(0.009)		
Log Markdown (Without CRS)	0.950	0.926	0.918
	(0.011)	(0.006)	(0.005)
NAICS2 $\times$ Year FE	No	Yes	Yes
Firm FE	No	No	Yes
Observations	69,500	69,500	69,500

Notes: This table reports the regression results from regressing the markdown estimate with the CRS restriction onto the markdown estimate without the CRS restriction. Columns (1) to (3) report the result without any fixed effects, with NAICS2 $\times$ year fixed effects, and with both NAICS2 $\times$ year and firm fixed effects. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A7: Price Markups and Firm Characteristics (CRS)

	Log Price Markup						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Wage Markdown	-0.104						
	(0.004)						
Log TFPR		0.002					
		(0.015)					
Log Labor Productivity			-0.045				
			(0.004)				
Log Sales				-0.016			
				(0.001)			
Log Wage Bill					-0.001		
					(0.001)		
Profit Share						0.007	
						(0.018)	
Log Labor Share VA							0.044
							(0.004)
NAICS2 FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NAICS2 $\times$ Year FE	No	No	No	No	No	No	No
Firm FE	No	No	No	No	No	No	No
Observations	69,500	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log price markups onto firm characteristics with NAICS2 fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Column (1) examines log wage markdowns, Columns (2) and (3) analyze log TFPR and log labor productivity, Columns (4) and (5) assess log sales and log wage bill, and Columns (6) and (7) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A8: Price Markups and Firm Characteristics (CRS)

	Log Price Markup						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Wage Markdown	-0.137						
	(0.005)						
Log TFPR		0.039					
		(0.020)					
Log Labor Productivity			-0.053				
			(0.004)				
Log Sales				-0.046			
				(0.002)			
Log Wage Bill					0.030		
					(0.003)		
Profit Share						-0.132	
						(0.011)	
Log Labor Share VA							0.039
							(0.004)
NAICS2 FE	No	No	No	No	No	No	No
NAICS2 $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	69,500	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log price markups onto firm characteristics with NAICS2 $\times$ year and firm fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Column (1) examines log wage markdowns, Columns (2) and (3) analyze log TFPR and log labor productivity, Columns (4) and (5) assess log sales and log wage bill, and Columns (6) and (7) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A9: Wage Markdowns and Firm Characteristics (CRS)

	Log Wage Markdown						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Price Markup	-2.627						
	(0.057)						
Log TFPR		1.205					
		(0.128)					
Log Labor Productivity			0.236				
			(0.018)				
Log Sales				0.206			
				(0.008)			
Log Wage Bill					0.133		
					(0.007)		
Profit Share						0.699	
						(0.097)	
Log Labor Share VA							-0.423
							(0.026)
NAICS2 FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NAICS2 $\times$ Year FE	No	No	No	No	No	No	No
Firm FE	No	No	No	No	No	No	No
Observations	69,500	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log wage markdowns onto firm characteristics with NAICS2 fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Column (1) examines log price markups, Columns (2) and (3) analyze log TFPR and log labor productivity, Columns (4) and (5) assess log sales and log wage bill, and Columns (6) and (7) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A10: Wage Markdowns and Firm Characteristics (CRS)

	Log Wage Markdown						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Price Markup	-2.283						
	(0.049)						
Log TFPR		0.931					
		(0.075)					
Log Labor Productivity			0.087				
			(0.014)				
Log Sales				0.240			
				(0.012)			
Log Wage Bill					-0.099		
					(0.011)		
Profit Share						0.106	
						(0.045)	
Log Labor Share VA							-0.190
							(0.016)
NAICS2 FE	No	No	No	No	No	No	No
NAICS2 $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	69,500	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log wage markdowns onto firm characteristics with NAICS2 $\times$ year and firm fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Column (1) examines log price markups, Columns (2) and (3) analyze log TFPR and log labor productivity, Columns (4) and (5) assess log sales and log wage bill, and Columns (6) and (7) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A11: Total Wedge and Firm Characteristics (CRS)

	Log Total Wedge					
	(1)	(2)	(3)	(4)	(5)	(6)
Log TFPR	1.207					
	(0.119)					
Log Labor Productivity		0.191				
		(0.018)				
Log Sales			0.189			
			(0.008)			
Log Wage Bill				0.132		
				(0.007)		
Profit Share					0.706	
					(0.093)	
Log Labor Share VA						-0.379
						(0.024)
NAICS2 FE	Yes	Yes	Yes	Yes	Yes	Yes
NAICS2 $\times$ Year FE	No	No	No	No	No	No
Firm FE	No	No	No	No	No	No
Observations	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log total wedges onto firm characteristics with NAICS2 fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Columns (1) and (2) analyze log TFPR and log labor productivity, Columns (3) and (4) assess log sales and log wage bill, and Columns (5) and (6) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

Table A12: Total Wedge and Firm Characteristics (CRS)

	Log Total Wedge					
	(1)	(2)	(3)	(4)	(5)	(6)
Log TFPR	0.970					
	(0.063)					
Log Labor Productivity		0.034				
		(0.014)				
Log Sales			0.195			
			(0.011)			
Log Wage Bill				-0.068		
				(0.010)		
Profit Share					-0.027	
					(0.039)	
Log Labor Share VA						-0.151
						(0.015)
NAICS2 FE	No	No	No	No	No	No
NAICS2 $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	69,500	69,500	69,500	69,500	69,500	69,500

Notes: This table reports the results from regressing log total wedges onto firm characteristics with NAICS2 $\times$ year and firm fixed effects. This table shows the results for markups and markdowns estimated with the CRS restriction. Columns (1) and (2) analyze log TFPR and log labor productivity, Columns (3) and (4) assess log sales and log wage bill, and Columns (5) and (6) evaluate profit share and labor share of value-added. Standard errors are reported in parentheses and are two-way clustered by firm and year. All figures are rounded in accordance with U.S. Census disclosure requirements.

B Additional Data Information

In this section, we discuss in more detail how the final dataset and key variables are constructed. In addition to the LBD and CRSP/Compustat Merged database, we utilize the following publicly available data: BEA NIPA Table 1.1.9 (Bureau of Economic Analysis, 2024), Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity (Board of Governors of the Federal Reserve System, 2024), and the CPI (Bureau of Labor Statistics, 2024). We use Lines 1 and 9 of BEA NIPA Table 1.1.9 to deflate financial statement line items, which correspond to the GDP implicit price deflator and non-residential fixed investment implicit price deflator, respectively. Line 9 is used for physical capital only. The 1-year U.S. Treasury data is used to proxy a one-year nominal risk-free rate and it is deflated by the CPI. We use the annual versions of all these datasets or convert them into annual series through averaging when we obtain them via FRED. The FRED series IDs are provided below.

1. BEA NIPA Table 1.1.9 Line 1 – FRED Series ID: A191RD3A086NBEA
2. BEA NIPA Table 1.1.9 Line 2 – FRED Series ID: A008RD3A086NBEA
3. Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity – FRED Series ID: DGS1
4. CPI – FRED Series ID: CPIAUCSL

B.1 Final Dataset Creation

Here we provide greater detail on how the final dataset is constructed given the raw data files. The final dataset is a merged CRSP/Compustat-LBD firm-level panel. We use the Compustat variable name where applicable. The steps are as follows:

1. Collapse the LBD files from the establishment level to the firm level using the firm-level identifiers
2. Merge the macroeconomic and time series datasets to the Compustat sample
3. Merge the LBD firm-level files to the Compustat sample
4. Replace the any missing values of XRD, XAD, XSGA, and XRENT with 0
5. Drop any observations without up to NAICS4 industry information
6. Generate intangible and physical capital investment and use forward iteration to generate the stock (See Appendix B.2 for information on how these are constructed)
7. Drop observations in NAICS2 industries 22, 52, 53, 92, and 99
8. Drop observations that are missing or non-positive in SALE, COGS, XSGA, AT, and the created capital stocks variables
9. Create the materials, profits, value-added, and labor productivity variables (See Appendix B.2 for information on how this is constructed)

10. Trim any observations that are in bottom and top 1% by year of COGS/SALE, XSGA/SALE, and materials/SALE. Also remove observations that are in the top 1% of XRD/SALE by year
11. Winsorize value-added, labor productivity, profit share, and investment rates at the 1% and 99% percentiles by year
12. Keep observations from 1976 onward
13. Run first-stage and second-stage estimation procedures of the production function estimation
14. Merge elasticity estimates to the current merged CRSP/Compustat-LBD panel
15. The firm-level elasticity and cost share estimates are winsorized at the 5% and 95% percentiles since these are used in the ratio estimators to reduce the impact of extreme values
16. Compute the markups, markdowns, and demand-shifter adjustment (See Appendices B.2 and C.3 for details on the variable construction and underlying theory, respectively)
17. Compute firm-level TFPR
18. Remove any observations with non-positive markups or markdowns

The final dataset is a firm-year panel that ranges from 1977 to 2019. The dataset starts in 1977 is because we required lagged inputs for production function estimation. This dataset is used in all analysis and is also the basis for various aggregations (industry- and aggregate-level).

B.2 Key Variable Construction

This section discusses how key variables are constructed that are not covered in the main text. It explains the construction of the production function inputs, including physical capital, intangible capital, and materials. Next, it describes the computation of firm value-added, labor productivity, labor share, and profit share. Then we discuss how to compute firm-level output elasticities and TFPR given the production function estimates. Additionally, the section details how the demand-shifter adjustment is constructed. Finally, the section explains the methodology for creating the aggregate indices used in the analysis (reported in Figures 1, 2, and A3 to A6) as well as the weighting approach applied in these calculations.

Production Function Inputs. The three remaining inputs (physical capital, intangible capital, and materials) are not directly taken from the existing data but rather are constructed using the existing data. Physical capital and intangible capital are both computed using a capitalization approach. Once these are computed, we can then compute materials.

For each firm's physical capital stock, we identify its first observation and initialize its value to the maximum of 0 and the first year's value of PPEGT (gross property, plant, and equipment). Then we take the first difference of PPENT (net property, plant, and equipment) to compute net investment and recursively define the subsequent year's physical capital stock. We deflate physical capital using the nonresidential fixed investment good deflator, which is line 9 of NIPA Table 1.1.9

(FRED ID: A008RD3A086NBEA). Since physical capital is recorded at the end of the year, we use the computed physical capital for year $t-1$ as the year t physical capital input in the estimation. As an example, in our notation, $k_{i,2000}$ refers to the end of 1999 end of year value of physical capital for firm i.

The firm's intangible capital stock is intialized following a similar procedure used by Eisfeldt and Papanikolaou (2013) and Peters and Taylor (2017). We define gross intangible investment as

$$\text{IntInv}_{i,t} = \text{XRD}_{i,t} + \text{XAD}_{i,t} + 0.3 \times (\text{XSGA}_{i,t} - \text{XRD}_{i,t} - \text{XAD}_{i,t}), \quad \text{(A1)}$$

where $\text{XRD}_{i,t}$ and $\text{XAD}_{i,t}$ are the research and development and advertising expenses from Compustat, respectively. We compute the sample median growth rate of intangible investment and initialize each firm's intangible capital stock as the maximum of 0 and the first year's intangible divided by the sum of the median growth rate plus depreciation. We set annual depreciation rate to be 30% following Eisfeldt and Papanikolaou (2013). We deflate these values using the GDP deflator which is from line 1 of NIPA Table 1.1.9 (FRED ID: A191RD3A086NBEA). As with physical capital, intangible capital follows the same timing convention.

Finally, recall that we define materials as $\text{COGS}_{i,t}$ plus $\text{XSGA}_{i,t}$ less the wage bill, rent (XRENT in Compustat), and the non-labor portion of the intangible investment. Since intangible investment is constructed using line items that are included in COGS or SGA, we must adjust for this in the definition of materials in Equation (8). Also, since we already remove the total wage bill, we only need to remove the non-labor portion of the intangible investment. Lehr (2023) estimates the labor share of research and development to be 79%. We use this figure and extend it for all intangible investment to compute the non-labor share of intangible investment.

Value-Added, Labor Productivity, Labor Share, and Profit. The firm's value-added is computed following Donangelo et al. (2019) and Seegmiller (2023), and is given below

$$\text{VA}_{i,t} = \text{OIBDP}_{i,t} + \Delta\text{INVFG}_{i,t} + \text{WB}_{i,t}, \quad \text{(A2)}$$

where $\text{OIBDP}_{i,t}$ is operating income before depreciation, $\Delta\text{INVFG}_{i,t}$ is the first difference in inventories, and $\text{WB}_{i,t}$ is the wage bill. The changes in inventories are set to 0 when missing. Given Equation (A2) we can compute log labor productivity and the labor share of value-added, which are given by Equations (A3) and (A4), respectively,

$$\ln(\text{Labor Productivity}_{i,t}) = \ln\left(\frac{\text{VA}_{i,t}}{\text{EMPLBD}_{i,t}}\right), \quad \text{(A3)}$$

$$\text{LSVA}_{i,t} = \frac{\text{WB}_{i,t}}{\text{VA}_{i,t}}, \quad \text{(A4)}$$

where $\text{EMPLBD}_{i,t}$ is the LBD-based employment measure.

The firm's profit in the data is defined as

$$\Pi_{i,t} = \text{OIBDP}_{i,t} - (r_{f,t} + \delta + \text{RP})K_{i,t}, \tag{A5}$$

where $r_{f,t}$ is the real risk-free rate (1-year Treasury rate), $\delta = 0.1$ is the annual depreciation rate of physical capital, $\text{RP} = 0.02$ a risk premium, $K_{i,t}$ is the physical capital stock in real terms. Equation (A5) follows the definition and imputations of De Loecker, Eeckhout and Unger (2020) for comparability. The firm's profit share is defined as the ratio of profit to sales.

Firm-Level Output Elasticities and TFPR. We estimate translog production functions, which is given by Equation (7). Therefore, the output elasticity of input j is given by

$$\theta^j_{i,t} = \beta_j + 2\beta_{j,j}x^j_{i,t} + \sum_{j' \in \mathcal{J}\setminus\{j\}} \beta_{j,j'}x^{j'}_{i,t}. \tag{A6}$$

We assume that $\beta_{j,j'} = \beta_{j',j}$ for all $j, j' \in \mathcal{J}$. With Equation (A6) we can compute the ratio estimators in Equations (4) and (5). Given the estimates of $\boldsymbol{\beta}$, firm-level log inputs $\mathbf{x}_{i,t}$, log output $y_{i,t}$, and first-stage estimate of $\varepsilon_{i,t}$, we can also recover log TFPR $\omega_{i,t}$ from re-arranging Equation (6) to isolate for $\omega_{i,t}$.

Demand-Shifter Adjustment. The demand-shifter adjustment for wage markdowns is constructed following Lemma 4 in Appendix C.3. We assume that we only need to adjust the wage markdown but not the price markup for demand-shifters. We set the elasticity term $\psi^{Q,j}_{i,t} = 1$ given the results from Cavenaile and Roldan-Blanco (2021). We estimate the ratio of demand-shifting labor to production labor with the following

$$\frac{X^{D,l}_{i,t}}{X^{Q,l}_{i,t}} = \max\left\{1, \frac{(\text{XAD}_{i,t} + \text{XRD}_{i,t})\alpha^{D,l}_{i,t}}{\text{WB}_{i,t} - (\text{XAD}_{i,t} + \text{XRD}_{i,t})\alpha^{D,l}_{i,t}}\right\}, \tag{A7}$$

where $\alpha^{D,l}_{i,t}$ is the labor share of XRD and XAD. We set $\alpha^{D,l}_{i,t} = 0.79$ based on the estimates of Lehr (2023). We lower bound the estimate to 1 and also winsorize from above at the 95th percentile to reduce the impact of extreme values as the estimator is a ratio.

Aggregate Price Markup and Wage Markdown Indices. First we compute the NAICS2 industry-level aggregate price markups and wage markdowns following the given weighting methodology. Then we normalize each NAICS2 aggregate series to its 1977 value to 1. From here we aggregate across these by taking a weighted average using the corresponding industry weights (i.e. the sales-weighted index uses the total NAICS2 industry-level sales; the unweighted or simple index uses the observation count as the weight). This indexation approach creates a weighted-average growth rate and bypasses the common multiplicative bias at the NAICS2 level since production functions are estimated at the NAICS2 level. In practice we find that this method produces very similar results to simply normalizing the final aggregate series' 1977 value to 1.

C Proofs and Derivations for Empirical Procedures

This section presents the proofs for the lemmas discussed in Section 3. We also formally derive and prove various results that relax the assumptions in the estimation of markups and markdowns but were omitted in the main text. We provide a brief discussion of the underlying intuition and additional empirical results that are relevant to the extension. Appendix C.2 discusses an extension that relaxes Assumption 1 (no adjustment costs) and Appendix C.3 discusses an extension that relaxes Assumption 5 (input used for direct production only).

C.1 Proofs of Propositions 1 and 2

The proofs of Propositions 1 and 2 are relatively straightforward. Recall that we need to rewrite the expressions for price markups and wage markdowns into objects that are empirically observable or can be estimated. We do this by rearranging the firm's first-order conditions into the markup and markdown relations using output elasticities and cost shares.

Proof of Proposition 1. We start by taking the first-order condition of the problem (3) with respect to a flexible input. This is given by

$$\begin{aligned}\frac{\partial P_{i,t}(Y_{i,t})Y_{i,t}}{\partial X^f_{i,t}} - \frac{\partial W^f_{i,t}(X^f_{i,t})X^f_{i,t}}{\partial X^f_{i,t}} &= \frac{\partial P_{i,t}(Y_{i,t})}{\partial Y_{i,t}}\frac{\partial Y_{i,t}}{\partial X^j_{i,t}}Y_{i,t} + P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^f_{i,t}} - \bar{W}^f_{i,t} \\ &= 0\end{aligned}$$

We can rearrange this expression as follows

$$\left[\frac{\partial P_{i,t}(Y_{i,t})}{\partial Y_{i,t}}\frac{Y_{i,t}}{P_{i,t}(Y_{i,t})} + 1\right]^{-1} = \frac{\partial Y_{i,t}}{\partial X^f_{i,t}}\frac{P_{i,t}(Y_{i,t})}{W^f_{i,t}},$$

Notice that the first two terms of the product on the right-hand side is the markup that follows the definition from (1). The expression for marginal costs follows from the dual problem (cost minimization). The right-hand side can be further rearranged as follows

$$\begin{aligned}\frac{\partial Y_{i,t}}{\partial X^f_{i,t}}\frac{P_{i,t}(Y_{i,t})}{W^f_{i,t}} &= \frac{\partial Y_{i,t}}{\partial X^f_{i,t}}\frac{X^f_{i,t}}{Y_{i,t}}\frac{P_{i,t}(Y_{i,t})Y_{i,t}}{W^f_{i,t}X^f_{i,t}} \\ &= \frac{\theta^f_{i,t}}{\alpha^f_{i,t}} = \mu_{i,t},\end{aligned}$$

which is the relationship in (4). □

Proof of Proposition 2. We follow similar steps as in the proof of Proposition 1 to prove Proposition 2. The first-order condition with respect to an input $j \in \mathcal{L}$ is given by

$$\frac{\partial P_{i,t}(Y_{i,t})Y_{i,t}}{\partial X^j_{i,t}} - \frac{\partial W^j_{i,t}(X^j_{i,t})X^j_{i,t}}{\partial X^j_{i,t}} = \frac{\partial P_{i,t}(Y_{i,t})}{\partial Y_{i,t}}\frac{\partial Y_{i,t}}{\partial X^j_{i,t}}Y_{i,t} + P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} - \frac{\partial W^j_{i,t}(X^j_{i,t})}{\partial X^j_{i,t}}X^j_{i,t} - W^j_{i,t}(X^j_{i,t})$$
$$= 0.$$

This expression can be rearranged as follows

$$\left[\frac{\partial P_{i,t}(Y_{i,t})}{\partial Y_{i,t}}\frac{Y_{i,t}}{P_{i,t}(Y_{i,t})} + 1\right]P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} = \frac{\partial W^j_{i,t}(X^j_{i,t})}{\partial X^j_{i,t}}X^j_{i,t} + W^j_{i,t}(X^j_{i,t})$$
$$\frac{\mu^{-1}_{i,t}}{W^j_{i,t}(X^j_{i,t})}P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} = \left[\frac{\partial W^j_{i,t}(X^j_{i,t})}{\partial X^j_{i,t}}\frac{X^j_{i,t}}{W^j_{i,t}(X^j_{i,t})} + 1\right].$$

The left-hand side of the first line is the marginal revenue product with respect to $X^j_{i,t}$; note that with product market power the MRPL accounts for the change in price when selling one more unit. Then on the second line, the left-hand side is the marginal revenue product of input j to the price of input j, which is the definition of a markdown following (2). This expression can be further changed to yield

$$\frac{\mu^{-1}_{i,t}}{W^j_{i,t}(X^j_{i,t})}P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} = \frac{P_{i,t}(Y_{i,t})Y_{i,t}}{W^j_{i,t}(X^j_{i,t})X^j_{i,t}}\frac{\partial Y_{i,t}}{\partial X^j_{i,t}}\frac{X^j_{i,t}}{Y_{i,t}}\mu^{-1}_{i,t}$$
$$= \frac{\theta^j_{i,t}}{\alpha^j_{i,t}}\mu^{-1}_{i,t} = \nu^j_{i,t},$$

thus we recover the relation in (5). □

C.2 Ratio Estimators with Adjustment Costs

This section considers an extension to Propositions 1 and 2 that relaxes Assumption 1 (no adjustment costs). The extension largely follows that of Bond et al. (2021) and Yeh, Macaluso and Hershbein (2022). First we consider the ratio estimator for markups in which there are adjustment costs to the flexible input and then we consider the case for input markdowns. Many adjustment cost specifications depend on past input choices, thus making some input decisions dynamic. Thus, we also need to relax Assumption 3 (static choice) as well, but that does not introduce any new complications. The firm's new problem is recursively given by

$$
\begin{gathered}
V(\mathbf{X}_{i,t-1};\Omega_{i,t}) = \max_{\mathbb{X}_{i,t}\in\mathbb{R}_{++}} \Pi_{i,t} + \beta\,\mathbb{E}_t\left[V(\mathbf{X}_{i,t};\Omega_{i,t+1})\right] \\
\text{subject to} \\
\Pi_{i,t} = P_{i,t}(Y_{i,t})Y_{i,t} - \sum_{j=1}^{J} W^j_{i,t}(X^j_{i,t})\left(X^j_{i,t} + \Phi^j(X^j_{i,t}, X^j_{i,t-1})\right) \\
Y_{i,t} \leq F(\mathbf{X}_{i,t};\omega_{i,t}),
\end{gathered}
\tag{A8}
$$

where $\Omega_{i,t}$ is a vector of exogenous stochastic state variables, $\Phi^j(\cdot,\cdot)$ is the adjustment cost function for input j, and $\beta \in (0,1)$ is the discount factor. In this problem, firms take past input decisions as given and these are state variables since they determine the adjustment costs for today.

The presence of adjustment costs means that markups and markdowns in Equations (1) and (2), respectively, do not only reflect present market power. This implies that a correction must be made to the standard ratio estimators to properly estimate market power. Similar to Yeh, Macaluso and Hershbein (2022), we assume that the adjustment cost function follows a standard quadratic functional form (Hall, 2004; Cooper and Haltiwanger, 2006; Cooper, Haltiwanger and Willis, 2007; Bloom, 2009), i.e.

$$
\Phi^j(X_1, X_2) = \frac{\gamma_j}{2}\left(\frac{X_1 - X_2}{X_2}\right)^2 X_2,
$$

for some scale parameter $\gamma_j \in \mathbb{R}_+$. However, throughout the proof we keep the expressions as general as possible since many of these steps apply for any $\Phi^j(\cdot,\cdot)$ so long standard smoothness conditions are satisfied. We introduce the lemma below.

Lemma 1 (Price Markup Ratio Estimator with Adjustment Costs)**.** *Suppose that the inputs $f \in \mathcal{F}$ do not necessarily satisfy Assumptions 1 and 3 and the firm solves the program (A8). Then, the ratio estimator in (4) is equal to*

$$
\frac{\theta^f_{i,t}}{\alpha^f_{i,t}} = \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1}\mathcal{A}^f_{i,t}, \tag{A9}
$$

where $\varepsilon_{i,t}$ is the elasticity of product demand for firm i, $g^f_{i,t}$ is the net growth rate of input f in period t, $g^{w,f}_{i,t}$ is the net growth rate of the total expenditure on input f in period t, and

$$
\mathcal{A}^f_{i,t} \equiv \left(1 + \gamma_f g^f_{i,t} - \frac{\beta\gamma_f}{2}\,\mathbb{E}_t\left[h(g^{w,f}_{i,t+1}, g^f_{i,t+1})\right]\right),
$$

where

$$
h(g^{w,f}_{i,t+1}, g^f_{i,t+1}) \equiv (1+g^{w,f}_{i,t+1})g^f_{i,t+1}\left((1+g^f_{i,t+1})^{-1}+1\right).
$$

Thus, the corrected ratio estimator that recovers the product market power of the firm is given by

$$\tilde{\mu}_{i,t} \equiv \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1} = \frac{\theta^f_{i,t}}{\alpha^f_{i,t}\mathcal{A}^f_{i,t}}. \tag{A10}$$

Proof. Take the first-order condition of the problem (A8) with respect to $X^f_{i,t}$, which yields

$$\begin{aligned}
0 &= \frac{\partial P_{i,t}(Y_{i,t})Y_{i,t}}{\partial X^f_{i,t}} - \bar{W}^f_{i,t}\left(1+\Phi^f_1(X^f_{i,t},X^f_{i,t-1})\right) + \beta\,\mathbb{E}_t\left[\frac{\partial V(\mathbf{X}_{i,t};\Omega_{t+1})}{\partial X^f_{i,t}}\right] \\
0 &= \frac{\partial P_{i,t}(Y_{i,t})}{\partial Y_{i,t}}\frac{\partial Y_{i,t}}{\partial X^f_{i,t}}Y_{i,t} + P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^f_{i,t}} - \bar{W}^f_{i,t}\left(1+\Phi^f_1(X^f_{i,t},X^f_{i,t-1})\right) + \beta\,\mathbb{E}_t\left[\frac{\partial V(\mathbf{X}_{i,t};\Omega_{t+1})}{\partial X^f_{i,t}}\right] \\
0 &= \left(-\varepsilon^{-1}_{i,t}+1\right)\frac{\partial Y_{i,t}}{\partial X^f_{i,t}} - \frac{\bar{W}^f_{i,t}}{P_{i,t}(Y_{i,t})}\left(1+\Phi^f_1(X^f_{i,t},X^f_{i,t-1})\right) - \beta\,\mathbb{E}_t\left[\frac{\bar{W}^f_{i,t+1}}{P_{i,t}(Y_{i,t})}\Phi^f_2(X^f_{i,t+1},X^f_{i,t})\right] \\
\frac{\theta^f_{i,t}}{\alpha^f_{i,t}} &= \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1}\left(\left(1+\Phi^f_1(X^f_{i,t},X^f_{i,t-1})\right) + \beta\,\mathbb{E}_t\left[\frac{\bar{W}^f_{i,t+1}}{\bar{W}^f_{i,t}}\Phi^f_2(X^f_{i,t+1},X^f_{i,t})\right]\right).
\end{aligned}$$

The second line expands the first term using the product rule. The third line uses the definition of the inverse demand elasticity and collects terms while evaluating the partial derivatives for the other terms. The third line also divides both sides by $P_{i,t}(Y_{i,t})$. The value function's derivative follows from the envelope theorem. The final line rearranges the terms to isolate the marginal product of $X^f_{i,t}$ and rescales both sides by $\bar{W}^f_{i,t}/P_{i,t}(Y_{i,t})$ to recover the ratio estimator.

From this we can see that the ratio estimator (4) recovers a component of the markup related to product market power (the first term) and another term related to adjustment costs (the remainder). The partial derivatives given the functional form of $\Phi^j(\cdot,\cdot)$ are

$$\begin{aligned}
\Phi^f_1(X_1,X_2) &= \gamma_f\left(\frac{X_1-X_2}{X_2}\right), \\
\Phi^f_2(X_1,X_2) &= \frac{\gamma_f}{2}\left(\frac{(X_2-X_1)(X_2+X_1)}{X_2^2}\right).
\end{aligned}$$

Define the net growth rate of the input f in year t as $g^f_{i,t} \equiv (X^f_{i,t}-X^f_{i,t-1})/X^f_{i,t-1}$ and the net growth rate of the total cost of input f in year t as $g^{w,f}_{i,t} \equiv (W^f_{i,t}X^f_{i,t} - W^f_{i,t-1}X^f_{i,t-1})/(W^f_{i,t-1}X^f_{i,t-1})$. Substitute these into the expression obtained from the prior steps, we obtain

$$\begin{aligned}
\frac{\theta^f_{i,t}}{\alpha^f_{i,t}} &= \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1}\left(1+\gamma_f g^f_{i,t}+\beta\,\mathbb{E}_t\left[\frac{\bar{W}^f_{i,t+1}}{\bar{W}^f_{i,t}}\frac{\gamma_f}{2}\left(\frac{(X^f_{i,t}-X^f_{i,t+1})(X^f_{i,t}+X^f_{i,t+1})}{\left(X^f_{i,t}\right)^2}\right)\right]\right)\\
&= \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1}\left(1+\gamma_f g^f_{i,t}+\frac{\beta\gamma_f}{2}\,\mathbb{E}_t\left[\frac{\bar{W}^f_{i,t+1}X^f_{i,t+1}}{\bar{W}^f_{i,t}X^f_{i,t}}\left(\frac{(X^f_{i,t}-X^f_{i,t+1})(X^f_{i,t}+X^f_{i,t+1})}{X^f_{i,t}X^f_{i,t+1}}\right)\right]\right)\\
&= \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1}\left(1+\gamma_f g^f_{i,t}+\frac{\beta\gamma_f}{2}\,\mathbb{E}_t\left[(1+g^{w,f}_{i,t+1})\left(\frac{(X^f_{i,t}-X^f_{i,t+1})(X^f_{i,t}+X^f_{i,t+1})}{X^f_{i,t}X^f_{i,t+1}}\right)\right]\right)\\
&= \frac{\varepsilon_{i,t}}{\varepsilon_{i,t}-1}\left(1+\gamma_f g^f_{i,t}-\frac{\beta\gamma_f}{2}\,\mathbb{E}_t\left[(1+g^{w,f}_{i,t+1})g^f_{i,t+1}\left((1+g^f_{i,t+1})^{-1}+1\right)\right]\right).
\end{aligned}$$

With this we can recover the expression in Equation (A10). □

The adjustment term $\mathcal{A}^f_{i,t}$ contains either parameters (β or γ_f) or terms that can be estimated with the data (the growth rates). Similar to before, Lemma 1 leads to the input markdown counterpart, which we outline in Lemma 2.

Lemma 2 (Input Markdown Ratio Estimator with Adjustment Costs)**.** *Suppose that the inputs $j \in \mathcal{L}$ do not necessarily satisfy Assumptions 1 and 3 and the firm solves the program (A8). Also, suppose that the flexible input set $\mathcal{F}$ follows the same assumptions as in Lemma 1. Then the modified input markdown estimator is given by*

$$\tilde{\nu}^j_{i,t} \equiv \frac{\eta^j_{i,t}+1}{\eta^j_{i,t}} = \frac{\nu^{j,\prime}_{i,t}-\mathcal{C}^j_{i,t}}{1+\mathcal{B}^j_{i,t}}, \tag{A11}$$

where $\eta^j_{i,j}$ is the elasticity of supply of input j and

$$\begin{aligned}
\nu^{j,\prime}_{i,t} &\equiv \frac{\theta^j_{i,t}}{\alpha^j_{i,t}}\tilde{\mu}^{-1}_{i,t},\\
\mathcal{B}^j_{i,t} &\equiv \frac{\gamma_j}{2}\frac{\left(g^j_{i,t}\right)^2}{1+g^j_{i,t}},\\
\mathcal{C}^j_{i,t} &\equiv -\mathcal{B}^j_{i,t}+\gamma_j g^j_{i,t}-\frac{\beta\gamma_j}{2}\,\mathbb{E}_t\left[h(g^{w,j}_{i,t+1},g^j_{i,t+1})\right].
\end{aligned}$$

Proof. We proceed similarly to the proof of Lemma 1 and take the first-order condition with respect to $X^j_{i,t}$. First we rearrange the terms of the first partial derivative of revenues,

$$\begin{aligned}\frac{\partial P_{i,t}(Y_{i,t})Y_{i,t}}{\partial X^j_{i,t}} &= \frac{\partial P_{i,t}(Y_{i,t})}{\partial Y_{i,t}}\frac{\partial Y_{i,t}}{\partial X^j_{i,t}}Y_{i,t} + P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} \\ &= \left(-\varepsilon^{-1}_{i,t} + 1\right) P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} \\ &= \tilde{\mu}^{-1}_{i,t} P_{i,t}(Y_{i,t})\frac{\partial Y_{i,t}}{\partial X^j_{i,t}}.\end{aligned}$$

This term is the left-hand side of the first-order condition after isolating for the revenue term; this term is also the marginal revenue product of labor. Dividing this by the $W_{i,t}(X^j_{i,t})$ recovers an expression similar to the original markdown ratio estimator but using $\tilde{\mu}_{i,t}$ instead.

$$\tilde{\mu}^{-1}_{i,t}\frac{P_{i,t}(Y_{i,t})}{W^j_{i,t}(X^j_{i,t})}\frac{\partial Y_{i,t}}{\partial X^j_{i,t}} = \frac{\theta^j_{i,t}}{\alpha^j_{i,t}}\tilde{\mu}^{-1}_{i,t},$$

where $\alpha^j_{i,t}$ is the cost share of j but this only includes the direct expenditures on j and excludes the adjustment costs.

Now we rearrange the first partial derivative of the cost terms and continuation value (where $C_{i,t}$ is the total cost and continuation value). We write $W^j_{i,t}(X^j_{i,t})$ as $W^j_{i,t}$ to simplify the notation

$$\begin{aligned}\frac{\partial C_{i,t}}{\partial X^j_{i,t}} &= \frac{\partial W^j_{i,t}}{\partial X^j_{i,t}}\left(X^j_{i,t} + \Phi^j(X^j_{i,t}, X^j_{i,t-1})\right) + W^j_{i,t}\left(1 + \Phi^j_1(X^j_{i,t}, X^j_{i,t-1})\right) - \beta\,\mathbb{E}_t\left[\frac{\partial V(\mathbf{X}_{i,t};\Omega_t)}{\partial X^j_{i,t}}\right] \\ &= \left(\frac{1}{\eta^j_{i,t}} + 1\right) W^j_{i,t} + \frac{\Phi^j(X^j_{i,t}, X^j_{i,t-1})}{X^j_{i,t}\eta^j_{i,t}}W^j_{i,t} + W^j_{i,t}\Phi^j_1(X^j_{i,t}, X^j_{i,t-1}) - \beta\,\mathbb{E}_t\left[\frac{\partial V(\mathbf{X}_{i,t};\Omega_t)}{\partial X^j_{i,t}}\right],\end{aligned}$$

We use the definition of the inverse labor supply elasticity and collect terms to derive the wage markdown expression. We can now divide both sides by $W^j_{i,t}$ to equate this with the left-hand side from before.

$$\begin{aligned}
\frac{\theta^j_{i,t}}{\alpha^j_{i,t}}\tilde{\mu}^{-1}_{i,t} &= \left(\frac{1}{\eta^j_{i,t}}+1\right) + \frac{\Phi^j(X^j_{i,t},X^j_{i,t-1})}{X^j_{i,t}\eta^j_{i,t}} + \Phi^j_1(X^j_{i,t},X^j_{i,t-1}) + \beta\,\mathbb{E}_t\left[\frac{W^j_{i,t+1}}{W^j_{i,t}}\Phi^j_2(X^j_{i,t+1},X^j_{i,t})\right] \\
&= \left(\frac{1}{\eta^j_{i,t}}+1\right) + \frac{\Phi^j(X^j_{i,t},X^j_{i,t-1})}{X^j_{i,t}\eta^j_{i,t}} + \gamma_j g^j_{i,t} - \frac{\beta\gamma_j}{2}\,\mathbb{E}_t\left[h(g^{w,j}_{i,t+1},g^j_{i,t+1})\right] \\
&= \left(\frac{1}{\eta^j_{i,t}}+1\right) + \frac{1}{\eta^j_{i,t}}\frac{\gamma_j}{2}\frac{\left(g^j_{i,t}\right)^2}{1+g^j_{i,t}} + \gamma_j g^j_{i,t} - \frac{\beta\gamma_j}{2}\,\mathbb{E}_t\left[h(g^{w,j}_{i,t+1},g^j_{i,t+1})\right] \\
&= \left(\frac{1}{\eta^j_{i,t}}+1\right)\left(1+\frac{\gamma_j}{2}\frac{\left(g^j_{i,t}\right)^2}{1+g^j_{i,t}}\right) - \frac{\gamma_j}{2}\frac{\left(g^j_{i,t}\right)^2}{1+g^j_{i,t}} + \gamma_j g^j_{i,t} - \frac{\beta\gamma_j}{2}\,\mathbb{E}_t\left[h(g^{w,j}_{i,t+1},g^j_{i,t+1})\right].
\end{aligned}$$

Part of these above steps follows the same approach in the proof of Lemma 1. The second line uses the definition of $h(\cdot,\cdot)$ and net growth rates to simplify the expression. The third line uses the definition of a net growth rate to expression the second term in terms of net growth rates and parameters. The final line collects common terms for the elasticity term. Finally, we can isolate for the elasticity term to yield

$$\frac{\eta^j_{i,t}+1}{\eta^j_{i,t}} = \frac{\nu^{j,\prime}_{i,t} - \mathcal{C}^j_{i,t}}{1+\mathcal{B}^j_{i,t}}.$$

We have derived the expression in Equation (A11). □

As with $\mathcal{A}^f_{i,t}$, the terms $\mathcal{B}^j_{i,t}$ and $\mathcal{C}^j_{i,t}$ are functions of given parameters or terms that can be estimated. Therefore, we can implement these adjustments in our empirical setting for both markups and markdowns. We proceed to show a range of possible adjustments given a calibration.

Consider the following calibration in which $\beta = 0.98$, $\gamma_j = 2, \forall j \in \{1,\ldots,J\}$, and all growth rates are set to within 15% to -15% and are known with certainty. These numbers are similar (or more aggressive) to that of Yeh, Macaluso and Hershbein (2022), who also find that the adjustments are relatively small. We present the results in Table A13; the estimates only need to be adjusted by at most 1.5% in either direction. Hence we do not make this adjustment in our baseline analysis and specification.

C.3 Ratio Estimators with Demand Shifting

This section relaxes Assumption 5 (input used for direct production only) and allows for inputs to shift or influence demand directly. Similar to Appendix C.2, the approach follows that of Yeh, Macaluso and Hershbein (2022) and Bond et al. (2021). The firm's problem is once again static but the firm's product inverse demand function now has an explicit demand-shifting term. The firm

Table A13: Impact of Adjustment Costs on Markups and Markdowns Estimates

Case	$g_{i,t}^{f}$ (1)	$g_{i,t}^{w,f}$ (2)	$g_{i,t}^{l}$ (3)	$g_{i,t}^{w,l}$ (4)	$\mathcal{A}_{i,t}^{f}$ (5)	$\mathcal{B}_{i,t}^{l}$ (6)	$\mathcal{C}_{i,t}^{l}$ (7)	$\hat{\mu}_{i,t}$ (8)	$\tilde{\mu}_{i,t}$ (9)	$\hat{\nu}_{i,t}^{l}$ (10)	$\tilde{\nu}_{i,t}^{l}$ (11)
Case 1	0.10	0.10	0.10	0.15	0.994	0.009	-0.024	1.000	1.006	1.000	1.009
Case 2	0.05	0.10	0.07	0.10	0.995	0.005	-0.011	1.000	1.005	1.000	1.001
Case 3	0.00	0.00	0.00	0.00	1.000	0.000	0.000	1.000	1.000	1.000	1.000
Case 4	-0.05	-0.10	-0.07	-0.10	0.991	0.005	-0.017	1.000	1.010	1.000	1.002
Case 5	-0.10	-0.10	-0.10	-0.15	0.986	0.011	-0.035	1.000	1.014	1.000	1.010

Notes: This table shows the estimates of the price markup and wage markdown with and without accounting for adjustment costs. Columns (1) to (4) present the firm-level growth rates of the flexible input, the total expenditure of the flexible input, wage, and the wage bill, respectively. We assume that the period $t+1$ growth rate is the same as the period t growth rate and that there is no uncertainty. Columns (5) to (7) report the computed values of $\mathcal{A}_{i,t}^{f}$, $\mathcal{B}_{i,t}^{l}$, and $\mathcal{C}_{i,t}^{l}$, respectively. Columns (8) and (10) display the estimated price markup and wage markdown using the standard ratio estimators, which are always set to 1 as a normalization. Columns (9) and (11) provide the estimates of the markup and markdown, respectively, that account for adjustment costs. We set $\beta = 0.98$ and $\gamma_l = \gamma_f = 2$ across all cases.

solves the following

$$\max_{\mathbf{X}_{i,t}^{Q}, \mathbf{X}_{i,t}^{D} \in \mathbb{R}_{+,+}^{J}} \Pi_{i,t} = P_{i,t}(Y_{i,t}, D_{i,t})Y_{i,t} - \sum_{j=1}^{J} W_{i,t}^{j}(X_{i,t}^{j}) \left(X_{i,t}^{Q,j} + X_{i,t}^{D,j} \right),$$
$$\text{subject to} \tag{A12}$$
$$Y_{i,t} \leq F(\mathbf{X}_{i,t}^{Q}; \omega_{i,t}),$$
$$D_{i,t} \leq \mathcal{D}(\mathbf{X}_{i,t}^{D}),$$

where $\mathbf{X}_{i,t}^{Q} = (X_{i,t}^{Q,1}, \ldots, X_{i,t}^{Q,J})$ is the vector of inputs used for production, $\mathbf{X}_{i,t}^{D} = (X_{i,t}^{D,1}, \ldots, X_{i,t}^{D,J})$ is the vector of inputs used for demand shifting, $X_{i,t}^{j} = X_{i,t}^{Q,j} + X_{i,t}^{D,j}$ is the total quantity of input j, and $\mathcal{D}(\cdot)$ is twice continuously differentiable and strictly increasing in all of its arguments. It is also assumed that $P_{i,t}(\cdot,\cdot)$ is increasing with respect to $D_{i,t}$ and that what matters for the inverse supply function of input j, $W_{i,t}^{j}(\cdot)$, is independent of how the inputs are split across the uses. With the new firm's problem established, we proceed to the lemmas that introduce the adjusted estimators.

Lemma 3 (Price Markup Ratio Estimator with Demand Shifters)**.** *Suppose that the inputs in $f \in \mathcal{F}$ do not necessarily satisfy Assumption 5 and the firm solves the program (A12). The price markup is given by*

$$\mu_{i,t} = \frac{\theta_{i,t}^{Q,f}}{\alpha_{i,t}^{Q,f}}, \tag{A13}$$

where $\theta_{i,t}^{Q,f}$ is the elasticity of output with respect to $X_{i,t}^{Q,f}$ and $\alpha_{i,t}^{Q,j}$ is the cost share of $X_{i,t}^{Q,f}$. If the researcher *estimates the price markup with the totals ($\theta^f_{i,t}/\alpha^f_{i,t}$), then the estimate is given by*

$$\frac{\theta^f_{i,t}}{\alpha^f_{i,t}} = \mu_{i,t}\frac{\psi^{Q,f}_{i,t}}{1 + X^{D,f}_{i,t}/X^{Q,f}_{i,t}},$$

where $\psi^{Q,f}_{i,t}$ is the elasticity of $X^{Q,f}_{i,t}$ with respect to $X^f_{i,t}$. Rearranging the above expression produces the price markup estimator that accounts for demand shifters

$$\mu_{i,t} = \frac{\theta^f_{i,t}}{\alpha^f_{i,t}}\frac{1 + X^{D,f}_{i,t}/X^{Q,f}_{i,t}}{\psi^{Q,f}_{i,t}}. \tag{A14}$$

Proof. The steps to recover Equation (A13) are the same as before but with the first-order conditions being with respect to $X^{Q,f}_{i,t}$. Next, we start by expanding the definition of $\theta^f_{i,t}$,

$$\begin{aligned}\theta^f_{i,t} &= \frac{\partial Y_{i,t}}{\partial X^f_{i,t}}\frac{X^f_{i,t}}{Y_{i,t}} \\ &= \frac{\partial Y_{i,t}}{\partial X^{Q,f}_{i,t}}\frac{\partial X^{Q,f}_{i,t}}{\partial X^f_{i,t}}\frac{X^f_{i,t}}{Y_{i,t}} \\ &= \frac{\partial Y_{i,t}}{\partial X^{Q,f}_{i,t}}\frac{X^{Q,f}_{i,t}}{Y_{i,t}}\frac{\partial X^{Q,f}_{i,t}}{\partial X^f_{i,t}}\frac{X^f_{i,t}}{X^{Q,f}_{i,t}} \\ &= \theta^{Q,f}_{i,t}\psi^{Q,f}_{i,t},\end{aligned}$$

the second line uses the chain rule and last line uses the definition of an elasticity. Now we proceed with expanding Equation (A13)

$$\begin{aligned}\frac{\theta^{Q,f}_{i,t}}{\alpha^{Q,f}_{i,t}} &= \frac{\theta^f_{i,t}}{\alpha^f_{i,t}}\frac{\alpha^f_{i,t}}{\alpha^{Q,f}_{i,t}}\frac{1}{\psi^{Q,f}_{i,t}} \\ &= \frac{\theta^f_{i,t}}{\alpha^f_{i,t}}\frac{1 + X^{D,f}_{i,t}/X^{Q,f}_{i,t}}{\psi^{Q,f}_{i,t}}.\end{aligned}$$

The second line follows from using the prior relationship and from the fact that

$$
\begin{aligned}
\frac{\alpha_{i,t}^{f}}{\alpha_{i,t}^{Q,f}} &= \frac{\alpha_{i,t}^{Q,f} + \alpha_{i,t}^{D,f}}{\alpha_{i,t}^{Q,f}} \\
&= 1 + \frac{\bar{W}_{i,t}^{f} X_{i,t}^{D,f}}{\bar{W}_{i,t}^{f} X_{i,t}^{Q,f}} \\
&= 1 + \frac{X_{i,t}^{D,f}}{X_{i,t}^{Q,f}}.
\end{aligned}
$$

Thus, we have recovered Equation (A14). □

One complication with implementing this procedure is estimating or recovering $\psi_{i,t}^{Q,f}$. Under certain assumptions a value can be assigned, for example the case in which $\psi_{i,t}^{Q,f} = 1$ implies that the level of $X_{i,t}^{f}$ has no effect on the share of the input used for production. Also notice that $\psi_{i,t}^{Q,j} = 1$ when $X_{i,t}^{f}$ is only used for production. Generally, one would need to observe $X_{i,t}^{Q,f}$ and $X_{i,t}^{f}$ in order to estimate $\psi_{i,t}^{Q,f}$, which reduces the need for this adjustment as there can be other approaches used to account for demand shifting. However, in practice one can set bounds on $\psi_{i,t}^{Q,f}$ based on prior research. This is discussed in further detail after Lemma 4 for the case of wage markdowns. For markups, we assume that our procedure adequately purges any demand shifting inputs and thus we do not need to use this adjustment; we mainly show and derive Lemma 3 for completeness.

Lemma 4 (Input Markdown Ratio Estimator with Demand Shifters)**.** *Suppose that the inputs in $j \in \mathcal{L}$ do not necessarily satisfy Assumption 5 and the firm solves the program* (A12)*. Also, suppose that the set of flexible inputs $\mathcal{F}$ follow the same assumption as in Lemma 3. Then the adjusted markdown estimator that corrects for demand shifters is given by*

$$
\nu_{i,t}^{j} = \hat{\nu}_{i,t}^{j} \frac{\psi_{i,t}^{Q,f}}{1 + X_{i,t}^{D,f} / X_{i,t}^{Q,f}} \frac{1 + X_{i,t}^{D,j} / X_{i,t}^{Q,j}}{\psi_{i,t}^{Q,j}}, \tag{A15}
$$

where $\hat{\nu}_{i,t}^{j}$ is the ratio estimator from Proposition 2.

Proof. Similar to the proof for Lemma 3, recovering the new expression for the input markdown follows a similar set of steps as before and it is straightforward to show that

$$
\nu_{i,t}^{j} = \frac{\theta_{i,t}^{Q,j}}{\alpha_{i,t}^{Q,j}} \frac{\alpha_{i,t}^{Q,f}}{\theta_{i,t}^{Q,f}}.
$$

The relationship between $\theta_{i,t}^{Q,j}$ and $\theta_{i,t}^{j}$ is derived just as in Lemma 3, thus $\theta_{i,t}^{j} = \theta_{i,t}^{Q,j} \psi_{i,t}^{Q,j}$. The rest of the proof is as follows

$$\frac{\theta^{Q,j}_{i,t}}{\alpha^{Q,j}_{i,t}}\frac{\alpha^{Q,f}_{i,t}}{\theta^{Q,f}_{i,t}} = \frac{\theta^{j}_{i,t}}{\alpha^{j}_{i,t}}\frac{\alpha^{f}_{i,t}}{\theta^{f}_{i,t}}\frac{\theta^{Q,j}_{i,t}}{\theta^{j}_{i,t}}\frac{\alpha^{j}_{i,t}}{\alpha^{Q,j}_{i,t}}\frac{\theta^{f}_{i,t}}{\theta^{Q,f}_{i,t}}\frac{\alpha^{Q,f}_{i,t}}{\alpha^{f}_{i,t}}$$
$$= \hat{\nu}^{j}_{i,t}\frac{\psi^{Q,f}_{i,t}}{1+X^{D,f}_{i,t}/X^{Q,f}_{i,t}}\frac{1+X^{D,j}_{i,t}/X^{Q,j}_{i,t}}{\psi^{Q,j}_{i,t}},$$

where the last line follows from the definition of the original ratio estimator from Proposition 2 and the breakdown of $\alpha^{j}_{i,t}$. Thus, we recovered Equation (A15) and we have proved Lemma 4. □

In Compustat, we can observe research and development expenses (XRD) and advertising expenses (XAD). Most of these expenses consist of labor expenses. Thus, we do not implement the adjustment from Lemma 3 for price markups and we can simplify the adjustment for wage markdowns to

$$\frac{\theta^{Q,j}_{i,t}}{\alpha^{Q,j}_{i,t}}\frac{\alpha^{Q,f}_{i,t}}{\theta^{Q,f}_{i,t}} = \hat{\nu}^{j}_{i,t}\frac{1+X^{D,j}_{i,t}/X^{Q,j}_{i,t}}{\psi^{Q,j}_{i,t}}.$$

Under the assumptions of Lemma 4, we estimate the ratio $X^{D,j}_{i,t}/X^{Q,j}_{i,t}$ as follows

$$\frac{X^{D,j}_{i,t}}{X^{Q,j}_{i,t}} = \frac{X\tilde{R}D_{i,t}+X\tilde{A}D_{i,t}}{WB_{i,t}-(X\tilde{R}D_{i,t}+X\tilde{A}D_{i,t})}.$$

where the tilde represents the wage expense component of XRD and XAD. We assume that 79% of XRD and XAD are wage expenses following Lehr (2023).

Next, we estimate $\psi^{Q,j}_{i,t}$. Cavenaile and Roldan-Blanco (2021) empirically document that larger publicly-traded firms spend proportionally less on R&D and advertising. Thus, this suggests that $\psi^{Q,j}_{i,t} \geq 1$ when combined with our other assumptions.

D Production Function Estimation Implementation

This section discusses the details on the implementation of the GMM procedure to estimate the second stage. First, we discuss the moment conditions and specifically how we incorporate the constant returns to scale (CRS) restriction. Then we discuss the procedure we utilize to address a common numerical optimization issue.

We implement the CRS restriction similar to the method described by Yeh, Macaluso and Hershbein (2022). However, first we discuss how the other moment conditions are constructed. Following the notation from Section 3.2 (in logs), we have materials $m_{i,t}$, physical capital $k_{i,t}$, labor $l_{i,t}$, and intangible capital $n_{i,t}$. We write the vector of the production function's parameters as

$$\boldsymbol{\beta} = (\beta_l, \beta_m, \beta_k, \beta_n, \beta_{l,l}, \beta_{l,m}, \beta_{l,k}, \beta_{l,n}, \beta_{m,m}, \beta_{m,k}, \beta_{m,n}, \beta_{k,k}, \beta_{k,n}, \beta_{n,n})^{\intercal}. \tag{A16}$$

Next, we define the vector of instruments and then we construct the first set of moment conditions. The vector of instruments is given by

$$\tilde{\mathbf{z}}_{i,t} = \left(\tilde{\mathbf{z}}_{1,i,t}^{\mathsf{T}}, \tilde{\mathbf{z}}_{2,i,t}^{\mathsf{T}}\right)^{\mathsf{T}}, \tag{A17}$$

where

$$\tilde{\mathbf{z}}_{1,i,t} = \left(l_{i,t-1},\, m_{i,t-1},\, k_{i,t},\, n_{i,t}\right)^{\mathsf{T}}, \tag{A18}$$

$$\tilde{\mathbf{z}}_{2,i,t} = \Big(l_{i,t-1}^2,\, l_{i,t-1}m_{i,t-1},\, l_{i,t-1}k_{i,t},\, l_{i,t-1}n_{i,t},\, m_{i,t-1}^2,\, m_{i,t-1}k_{i,t},\, m_{i,t-1}n_{i,t},\, k_{i,t}^2,\, k_{i,t}n_{i,t},\, n_{i,t}^2\Big)^{\mathsf{T}}. \tag{A19}$$

Notice that in Equations (A17) to (A19) we impose a timing assumption for identification. Following the standard assumptions, we assume that the firm observes their idiosyncratic productivity shock $\xi_{i,t}$ and then make their input decisions. However, since the current stock of physical and intangible capital are assumed to be chosen in the period before, their current value is orthogonal to $\xi_{i,t}$. Similarly, since materials and labor are chosen contemporaneously, we use their lagged values as instruments as those are uncorrelated with $\xi_{i,t}$. Now we move on to the CRS restriction's moment condition.

Let $\Sigma_{i,t}(\cdot)$ denote a firm's returns to scale that takes the production function parameters as an input. The returns to scale are given by

$$\Sigma_{i,t}(\boldsymbol{\beta}) = \sum_{j\in\mathcal{J}} \frac{\partial f(\mathbf{x}_{i,t};\boldsymbol{\beta})}{\partial x_{i,t}^j}. \tag{A20}$$

CRS implies that Equation (A20) is set to 1. Therefore, the moment conditions that includes the CRS restriction are given by

$$\mathbb{E}\begin{bmatrix}\xi_{i,t}(\hat{\boldsymbol{\beta}})\cdot\tilde{\mathbf{z}}_{i,t}\\ \Sigma_{i,t}(\hat{\boldsymbol{\beta}})-1\end{bmatrix} = \mathbf{0}. \tag{A21}$$

Since we are using a translog production function, we can write the CRS restriction as a linear operator. Let $\tilde{\boldsymbol{\beta}} = (1,\, \boldsymbol{\beta}^{\mathsf{T}})^{\mathsf{T}}$ and $\tilde{\mathbf{x}}_{i,t} = \left(1,\, \mathbf{x}_{i,t}^{\mathsf{T}}\right)^{\mathsf{T}}$, then we have

$$\Sigma_{i,t}(\boldsymbol{\beta}) - 1 = (R\tilde{\boldsymbol{\beta}})^{\mathsf{T}}\tilde{\mathbf{x}}_{i,t}, \tag{A22}$$

where

$$R = \begin{pmatrix} -1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 2 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix}.$$

When we compute the norm of the moment conditions, we use the identity matrix as the weight matrix for our baseline specification with the CRS restriction. To estimate a version without this restriction, we replace the last diagonal entry of the weight matrix from 1 to 0, which corresponds to Equation (A22).

Once we construct these objects, we simply need to run the optimization procedure to recover the estimates. However, optimization routines do not necessarily find the global minimum for a given starting point and boundary restriction. Indeed, if we utilize different starting points we generally obtain different solutions. Therefore, we run the estimation 1,000 times with different (randomly selected) starting points within the boundaries. We collect the estimates and the value of the objective function for each iteration across all industries. We select the estimates associated with the minimum evaluated objective function value across all iterations for each industry to be the "global" solution. While this approach still does not guarantee a global solution, we find that the approach is robust to picking 100, 200, and 500 iterations. The parameter bounds are set to be $[0, 1]$ for the first-order terms and $[-0.05, 0.05]$ for the second-order and interaction terms. The estimated parameters for all industries have all their respective estimates within these bounds.

E Proofs and Derivations for the Model

E.1 Derivation of the Ideal Price and Wage Indices

We formally derive the ideal price and wage indices (P_t^I and W_t^I) and show how they are linked to the competition indices (P_t and W_t). We begin with defining the ideal price and wage indices from the household's problem. We drop time subscripts for ease of notation. The ideal price index P^I is defined as follows

$$P^I = \frac{M}{V_P(\mathbf{p}, M)}, \tag{A23}$$

where M is the total income and $V_P(\mathbf{p}, M)$ is the indirect utility function which is given by

$$V_P(\mathbf{p}, M) \equiv \max_{\mathbf{x}} \{u(\mathbf{x}) \mid \mathbf{p} \cdot \mathbf{x} \leq M\},$$

where $\mathbf{x}$ is the vector of consumption goods. The ideal wage index is derived similarly, however, it is related to the disutility minimization problem for a household that wants to meet a certain budget or earnings level. Therefore, the ideal wage index is given by

$$W^I = \frac{E}{V_W(\mathbf{w}, E)}, \tag{A24}$$

where E is the target earnings and $V_W(\mathbf{w}, E)$ is the indirect disutility function which is given by

$$V_W(\mathbf{w}, E) \equiv \min_{\mathbf{l}} \{v(\mathbf{l}) \mid \mathbf{w} \cdot \mathbf{l} \geq E\},$$

where $v(\cdot)$ is the disutility function over labor and l is the vector of labor supplied. Since the indirect utility and disutility are homogeneous of degree 0, P^I and W^I are homogeneous of degree 1 with respect to $\mathbf{p}$ and $\mathbf{w}$, respectively, given the functional forms of Equations (A23) and (A24). Furthermore, the ideal indices only depend on the prices/wages. From here, it is helpful to denote $P^I = P^I(\mathbf{p})$ and $W^I = W^I(\mathbf{w})$.

We differentiate the indirect utility and disutility functions with respect to their arguments and apply Roy's Identity to recover the demand and supply functions

$$\begin{aligned}
\frac{\partial V_P(\mathbf{p}, M)}{\partial M} &= \frac{1}{P^I(\mathbf{p})}, \\
\frac{\partial V_P(\mathbf{p}, M)}{\partial p_i} &= -\frac{M}{P^I(\mathbf{p})^2}\frac{\partial P^I(\mathbf{p})}{\partial p_i}, \\
\frac{\partial V_W(\mathbf{w}, E)}{\partial E} &= \frac{1}{W^I(\mathbf{w})}, \\
\frac{\partial V_W(\mathbf{w}, E)}{\partial w_i} &= -\frac{E}{W^I(\mathbf{w})^2}\frac{\partial W^I(\mathbf{w})}{\partial w_i}.
\end{aligned}$$

Then with Roy's Identity it follows that

$$\begin{aligned}
x_i &= \frac{M}{P^I(\mathbf{p})}\frac{\partial P^I(\mathbf{p})}{\partial p_i}, \\
l_i &= \frac{E}{W^I(\mathbf{w})}\frac{\partial W^I(\mathbf{w})}{\partial w_i}.
\end{aligned}$$

If we multiple the above expressions by p_i and w_i, respectively, and apply the elasticity definition we recover

$$\frac{\partial \ln P^I(\mathbf{p})}{\partial \ln p_i} = r_i\left(\frac{p_i}{A(\mathbf{p})}\right), \tag{A25}$$

$$\frac{\partial \ln W^I(\mathbf{w})}{\partial \ln w_i} = s_i\left(\frac{w_i}{B(\mathbf{w})}\right). \tag{A26}$$

Recall that we define $A(\mathbf{p}) \equiv P$ and $B(\mathbf{w}) \equiv W$.

Finally, we can relate the ideal indices to their competition index counterparts. It useful to rearrange to Equations (A25) and (A26) as follows

$$\begin{aligned}
\frac{\partial P^I(\mathbf{p})}{\partial p_i}\frac{1}{P^I(\mathbf{p})} &= \frac{r_i(a_i)}{p_i}, \\
\frac{\partial W^I(\mathbf{w})}{\partial w_i}\frac{1}{W^I(\mathbf{w})} &= \frac{s_i(b_i)}{w_i}.
\end{aligned}$$

We use the same notation as in the main text for a_i and b_i. Since the ideal price indices are homogeneous of degree 1, we can rewrite the above as

$$\frac{\partial P^I(\mathbf{a})}{\partial a_i}\frac{1}{P^I(\mathbf{a})A(\mathbf{p})} = \frac{r_i(a_i)}{p_i} \implies \frac{\partial P^I(\mathbf{a})}{\partial a_i}\frac{1}{P^I(\mathbf{a})} = \frac{r_i(a_i)}{a_i},$$
$$\frac{\partial W^I(\mathbf{b})}{\partial b_i}\frac{1}{W^I(\mathbf{b})B(\mathbf{w})} = \frac{s_i(b_i)}{w_i} \implies \frac{\partial W^I(\mathbf{b})}{\partial b_i}\frac{1}{W^I(\mathbf{b})} = \frac{s_i(b_i)}{b_i},$$

where $\mathbf{a}$ and $\mathbf{b}$ are the vectors containing a_i and b_i, respectively. Integrate both sides with respect to a_i and b_i, respectively, and by the firm index i to obtain

$$\ln P^I(\mathbf{a}) = \int_{i=0}^{1}\int_{z=c_P}^{a_i}\frac{r_i(z)}{z}\,dz\,di,$$
$$\ln W^I(\mathbf{b}) = \int_{i=0}^{1}\int_{z=c_W}^{b_i}\frac{s_i(z)}{z}\,dz\,di,$$

where c_P and c_W are constants. Since the ideal indices are homogeneous of degree 1, this results in

$$\ln P^I(\mathbf{p}) = \ln A(\mathbf{p}) + \int_{i=0}^{1}\int_{z=c_P}^{a_i}\frac{r_i(z)}{z}\,dz\,di, \tag{A27}$$
$$\ln W^I(\mathbf{w}) = \ln B(\mathbf{w}) + \int_{i=0}^{1}\int_{z=c_W}^{b_i}\frac{s_i(z)}{z}\,dz\,di. \tag{A28}$$

Thus, we have the explicit relationship between the ideal and competition indices. Matsuyama and Ushchev (2017) and Matsuyama (2023) provide a more extensive treatment of these derivations as well as a more thorough discussion of the properties of HSA aggregator.

E.2 Derivation of the Firm's First-Order Conditions and Elasticities

We derive the firm's first-order conditions and elasticities for the model in this section. Let $\lambda_{i,t}$ be the multiplier on the production function constraint and substitute in the demand and supply functions into the objective function. Then the first-order condition with respect to $p_{i,t}$ is given by

$$\begin{aligned}\frac{\partial p_{i,t}y_{i,t}}{\partial p_{i,t}} - \lambda_{i,t}\frac{\partial y_{i,t}}{\partial p_{i,t}} &= \frac{r_i'(a_{i,t})}{P_t}P_t^I Y_t - \lambda_{i,t}\frac{r_i'(a_{i,t})p_{i,t}P_t^{-1}P_t^I Y_t - r_i(a_{i,t})P_t^I Y_t}{p_{i,t}^2} \\ &= r_i'(a_{i,t}) - \lambda_{i,t}\frac{(r_i'(a_{i,t})a_{i,t} - r_i(a_{i,t}))}{p_{i,t}a_{i,t}} = 0.\end{aligned}$$

The first line follows from the constraint imposed by demand and the second line removes excess terms by taking advantage of the fact that the expression is equal to 0. Now the take the first-order condition with respect to $w_{i,t}$, which follows a similar series of steps,

$$
\begin{aligned}
-\frac{\partial w_{i,t} l_{i,t}}{\partial w_{i,t}} + \lambda_{i,t}\frac{\partial l_{i,t}}{\partial w_{i,t}} &= -\frac{s_i'(b_{i,t})}{W_t} W_t^I L_t + \lambda_{i,t}\omega_i \frac{s_i'(b_{i,t}) b_{i,t} W_t^{-1} W_t^I L_t - s_i(b_{i,t}) W_t^I L_t}{w_{i,t}^2} \\
&= -s_i'(b_{i,t}) + \lambda_{i,t}\omega_i \frac{\left(s_i'(b_{i,t}) b_{i,t} - s_i(b_{i,t})\right)}{w_{i,t} b_{i,t}} = 0.
\end{aligned}
$$

Rearrange both expressions to isolate for $\lambda_{i,t}$, then we obtain

$$
\begin{aligned}
\frac{p_{i,t} r_i'(a_{i,t}) a_{i,t}}{r_i'(a_{i,t}) a_{i,t} - r_i(a_{i,t})} &= \frac{\omega_i^{-1} w_{i,t} s_i'(b_{i,t}) b_{i,t}}{s_i'(b_{i,t}) b_{i,t} - s_i(b_{i,t})} \\
p_{i,t} &= \mu_i(a_{i,t}) \nu_i(b_{i,t}) \frac{w_{i,t}}{\omega_i}.
\end{aligned}
$$

The expressions for markups and markdowns are derived later in this section.

Now we derive Equations (25) and (26). Note that the derivations here are general to any standard production function and number of inputs. Starting from the definition, it follows that

$$
\begin{aligned}
-\frac{\partial \ln y_{i,t}}{\partial \ln p_{i,t}} &= -\frac{\partial y_{i,t}}{\partial p_{i,t}} \frac{p_{i,t}}{y_{i,y}} \\
&= -\frac{r_i'(a_{i,t}) p_{i,t} P_t^{-1} P_t^I Y_t - r_i(a_{i,t}) P_t^I Y_t}{p_{i,t}^2} \frac{p_{i,t}}{y_{i,t}} \\
&= -\frac{\left(r_i'(a_{i,t}) a_{i,t} - r_i(a_{i,t})\right) P_t^I Y_t}{p_{i,t} y_{i,t}} \\
&= 1 - \frac{r_i'(a_{i,t}) a_{i,t}}{r_i(a_{i,t})},
\end{aligned}
$$

where the second line follows from the quotient rule applied onto the definition of the residual demand function, the third line follows from the definition of $a_{i,t}$, and the final line follows from the definition of $r_i(a_{i,t})$. The derivation of $\eta_i(b_{i,t})$ follows a similar approach.

$$
\begin{aligned}
\frac{\partial \ln l_{i,t}}{\partial \ln w_{i,t}} &= \frac{\partial l_{i,t}}{\partial w_{i,t}} \frac{w_{i,t}}{l_{i,t}} \\
&= \frac{s_i'(b_{i,t}) w_{i,t} W_t^{-1} W_t^I L_t - s_i(b_{i,t}) W_t^I L_t}{w_{i,t}^2} \frac{w_{i,t}}{l_{i,t}} \\
&= \frac{\left(s_i'(b_{i,t}) - s_i(b_{i,t})\right) W_t^I L_t}{w_{i,t} l_{i,t}} \\
&= \frac{s_i'(b_{i,t})}{s_i(b_{i,t})} - 1.
\end{aligned}
$$

We can use the definitions of the price markup and wage markdown to obtain the following

$$\begin{aligned}
\mu_i(a_{i,t}) &= \frac{\varepsilon_i(a_{i,t})}{\varepsilon_i(a_{i,t})-1} \\
&= \left(1-\frac{r_i'(a_{i,t})a_{i,t}}{r_i(a_{i,t})}\right)\left(\frac{r_i'(a_{i,t})a_{i,t}}{r_i(a_{i,t})}\right)^{-1} \\
&= \frac{r_i(a_{i,t})-r_i'(a_{i,t})a_{i,t}}{r_i'(a_{i,t})a_{i,t}}, \\
\nu_i(b_{i,t}) &= \frac{\eta_i(b_{i,t})+1}{\eta_i(b_{i,t})} \\
&= \left(\frac{s_i'(b_{i,t})b_{i,t}}{s_i(b_{i,t})}\right)\left(\frac{s_i'(b_{i,t})b_{i,t}}{s_i(b_{i,t})}-1\right)^{-1} \\
&= \frac{s_i'(b_{i,t})b_{i,t}}{s_i'(b_{i,t})b_{i,t}-s_i(b_{i,t})}.
\end{aligned}$$

We can substitute these expressions back into the first-order conditions to recover the pricing and wage rules (Equations (29) and (30)).

E.3 Proof of Proposition 3

Proof of Proposition 3. Suppose that ω_i increases but the firm's decisions remain at the old equilibrium choices. Then the pricing rule (29) fails to hold at equality and is given by

$$p_{i,t} > \mu_i(a_{i,t})\nu_i(b_{i,t})\frac{w_{i,t}}{\omega_i}.$$

It is helpful to rearrange this to yield

$$p_{i,t}\mu_i(a_{i,t})^{-1} > \nu_i(b_{i,t})\frac{w_{i,t}}{\omega_i}.$$

In order for this condition to hold with equality, we require $p_{i,t}$ to decrease. This decreases the left-hand side because of Assumption 9. Given this, it is sufficient to show that the remaining terms on the right-hand side weakly increase. Since aggregate conditions are constant, $a_{i,t}$ decreases, which in turn implies that $b_{i,t}$. This also implies that $w_{i,t}$ increases.

Given these results, it is straightforward to show that revenues $p_{i,t}y_{i,t}$ and the wage bill $w_{i,t}l_{i,t}$ increase. First, differentiate revenues with respect to price $p_{i,t}$ which results in

$$\frac{\partial p_{i,t}y_{i,t}}{\partial p_{i,t}} = \frac{r_i'(a_{i,t})P_t^I Y_t}{P_t} < 0,$$

by the demand constraint and Assumption 7, thus a decrease in price leads to an increase in revenues. Similarly, differentiate the wage bill with respect to wages $w_{i,t}$

$$\frac{\partial w_{i,t} l_{i,t}}{\partial w_{i,t}} = \frac{s_i'(b_{i,t}) W_t^I L_t}{W_t} > 0,$$

by the supply constraint and Assumption 7, thus a decrease in price, which leads to an increase in wages, increases the wage bill.

Now we differentiate the price markup and wage markdown with respect to $a_{i,t}$ and $b_{i,t}$, respectively (which is equivalent to differentiating with respect to prices and wages up to a positive scalar).

$$\begin{aligned}
\frac{\partial \mu_i(a_{i,t})}{\partial a_{i,t}} &= \frac{\partial \mu_i(a_{i,t})}{\partial \varepsilon_i(a_{i,t})} \frac{\partial \varepsilon_i(a_{i,t})}{\partial a_{i,t}} \\
&= -\frac{1}{\left(\varepsilon_i(a_{i,t}) - 1\right)^2} \frac{\partial \varepsilon_i(a_{i,t})}{\partial a_{i,t}} > 0, \\
\frac{\partial \nu_i(b_{i,t})}{\partial b_{i,t}} &= \frac{\partial \nu_i(a_{i,t})}{\partial \eta_i(b_{i,t})} \frac{\partial \eta_i(b_{i,t})}{\partial b_{i,t}} \\
&= -\frac{1}{\eta_i(b_{i,t})^2} \frac{\partial \eta_i(b_{i,t})}{\partial b_{i,t}} > 0.
\end{aligned}$$

The inequalities follow from Assumptions 7, 8, and 10 as $\varepsilon_i(a_{i,t}) > 1$ and $\eta_i(b_{i,t}) > 0$ and the elasticities are strictly decreasing with respect to their relative price/wage. Since $a_{i,t}$ decreases and $b_{i,t}$ increases, markups decrease and markdowns increase. Thus, we have proved the required results. □

E.4 Proof of Proposition 4

Before we prove Proposition 4, it is useful to establish some intermediate results via Lemmas 5 and 6. Lemma 5 relates pass-throughs ($\rho_{i,t}^{p}$ and $\rho_{i,t}^{w}$) to markups and markdowns (as well as their derivatives). Lemma 6 provides a link between the effective price pass-through $\rho_{i,t}^{p,\omega}$ and effective wage pass-through $\rho_{i,t}^{w,\omega}$. Finally, it is also helpful to define the following

$$\begin{aligned}
f(a_{i,t}) &= \frac{s(b_{i,t})}{b_{i,t}} \frac{W_t^I L_t}{W_t} = l_{i,t}, \\
h(b_{i,t}) &= \frac{r(a_{i,t})}{a_{i,t}} \frac{P_t^I Y_t}{P_t} = y_{i,t}.
\end{aligned}$$

Lemma 5. *The price and wage pass-throughs can be expressed as*

$$\rho_{i,t}^{p} = \left(1 - \frac{a_{i,t}\mu_i'(a_{i,t})}{\mu_i(a_{i,t})}\right)^{-1}, \tag{A29}$$

$$\rho_{i,t}^{w} = \left(1 + \frac{b_{i,t}\nu_i'(b_{i,t})}{\nu_i(b_{i,t})}\right)^{-1}. \tag{A30}$$

Proof. For Equation (A29) we expand the definition

$$\begin{aligned}
\rho_{i,t}^{p} &= \frac{\partial \ln a_{i,t}}{\partial \ln \mathrm{MC}_{i,t}} \\
&= \frac{\partial \ln \left(\mu_i(a_{i,t})\mathrm{MC}_{i,t}\right)}{\partial \ln \mathrm{MC}_{i,t}} \\
&= \frac{\partial \ln \mu_i(a_{i,t})}{\partial \ln \mathrm{MC}_{i,t}} + 1 \\
&= \frac{\partial \ln \mu_i(a_{i,t})}{\partial \ln a_{i,t}} \frac{\partial \ln a_{i,t}}{\ln \mathrm{MC}_{i,t}} + 1 \\
&= \left(1 - \frac{a_{i,t}\mu_i'(a_{i,t})}{\mu_i(a_{i,t})}\right)^{-1}.
\end{aligned}$$

We use a similar approach for Equation (A30)

$$\begin{aligned}
\rho_{i,t}^{w} &= \frac{\partial \ln b_{i,t}}{\partial \ln \mathrm{MRPL}_{i,t}} \\
&= \frac{\partial \ln \left(\nu_i(b_{i,t})^{-1}\mathrm{MRPL}_{i,t}\right)}{\partial \ln \mathrm{MRPL}_{i,t}} \\
&= \frac{\partial \ln \nu_i(b_{i,t})^{-1}}{\partial \ln \mathrm{MRPL}_{i,t}} + 1 \\
&= -\frac{\partial \ln \nu_i(b_{i,t})}{\partial \ln b_{i,t}} \frac{\partial \ln b_{i,t}}{\ln \mathrm{MRPL}_{i,t}} + 1 \\
&= \left(1 + \frac{b_{i,t}\nu_i'(b_{i,t})}{\nu_i(b_{i,t})}\right)^{-1}.
\end{aligned}$$

From these we also can derive

$$\begin{aligned}
\frac{\partial \ln \mu_i(a_{i,t})}{\partial \ln a_{i,t}} &= 1 - \frac{1}{\rho_{i,t}^{p}}, \\
\frac{\partial \ln \nu_i(b_{i,t})}{\partial \ln b_{i,t}} &= \frac{1}{\rho_{i,t}^{w}} - 1.
\end{aligned}$$

□

Lemma 6. *The effective price pass-through is related to the effective wage pass-through as follows:*

$$\rho_{i,t}^{p,\omega} = \frac{\rho_{i,t}^{p}}{\rho_{i,t}^{w}} \rho_{i,t}^{w,\omega} - \rho_{i,t}^{p}. \tag{A31}$$

Proof. We start with how price pass-through is related to the effective price pass-through which yields

$$\begin{aligned}
\rho_{i,t}^{p,\omega} &= \frac{\partial \ln a_{i,t}}{\partial \ln \mathrm{MC}_{i,t}} \frac{\partial \ln \mathrm{MC}_{i,t}}{\partial \ln \omega_i} \\
&= \rho_{i,t}^{p} \frac{\partial \ln v_i(b_{i,t}) w_{i,t} \omega_i^{-1}}{\partial \ln \omega_i} \\
&= \rho_{i,t}^{p} \left(\frac{\partial \ln v_i(b_{i,t})}{\partial \ln \omega_i} + \frac{\partial \ln w_{i,t}}{\partial \ln \omega_i} - 1 \right) \\
&= \rho_{i,t}^{p} \left(\frac{\partial \ln v_i(b_{i,t})}{\partial \ln b_{i,t}} \frac{\partial \ln b_{i,t}}{\partial \ln \omega_i} + \frac{\partial \ln b_{i,t}}{\partial \ln \omega_i} - 1 \right) \\
&= \rho_{i,t}^{p} \left(\left[\frac{1}{\rho_{i,t}^{w}} - 1 \right] \rho_{i,t}^{w,\omega} + \rho_{i,t}^{w,\omega} - 1 \right) \\
&= \frac{\rho_{i,t}^{p}}{\rho_{i,t}^{w}} \rho_{i,t}^{w,\omega} - \rho_{i,t}^{p}.
\end{aligned}$$

The second line follows the definition of marginal cost and the remaining lines employ the results from Lemma 5 or are standard algebra. Thus, we recover Equation (A31). □

With Lemmas 5 and 6, we proceed to prove Proposition 4.

Proof of Proposition 4. We start by expanding the expression for the effective wage pass-through

$\rho_{i,t}^{w,\omega}$ and substitute in the expressions using the the functions $h(\cdot)$ and $f(\cdot)$.

$$\begin{aligned}
\frac{\partial \ln b_{i,t}}{\partial \ln \omega_i} &= \frac{\ln f^{-1}\left(h(a_{i,t})\omega_{i,t}^{-1}\right)}{\partial \ln \omega_i} \\
&= \frac{1}{f'(b_{i,t})} \frac{\frac{\partial y_{i,t}}{\partial p_{i,t}}\frac{\partial p_{i,t}}{\partial \omega_i}\omega_i - y_{i,t}}{\omega_i^2} \frac{\omega_i}{b_{i,t}} \\
&= \frac{b_{i,t}^2}{s_i'(b_{i,t})b_{i,t} - s_i(b_{i,t})} \frac{W_t}{W_t^I L_t} \frac{-\varepsilon_i(a_{i,t})\frac{y_{i,t}}{a_{i,t}}\frac{\partial a_{i,t}}{\partial \omega_i}\omega_i - \omega_i l_{i,t}}{\omega_i^2} \frac{\omega_i}{b_{i,t}} \\
&= \frac{b_{i,t}}{\eta_i(b_{i,t}) l_{i,t}} \frac{-\varepsilon_i(a_{i,t})\frac{y_{i,t}}{a_{i,t}}\frac{\partial a_{i,t}}{\partial \omega_i}\omega_i - \omega_i l_{i,t}}{\omega_i^2} \frac{\omega_i}{b_{i,t}} \\
&= \frac{1}{\eta_i(b_{i,t})}\left(-\varepsilon_i(a_{i,t})\frac{\omega_i}{a_{i,t}}\frac{\partial a_{i,t}}{\partial \omega_i} - 1\right) \\
&= -\frac{1}{\eta_i(b_{i,t})}\left(\varepsilon_i(a_{i,t})\frac{\partial \ln a_{i,t}}{\partial \ln \omega_i} + 1\right).
\end{aligned}$$

The second line uses the fact that $h(a_{i,t}) = y_{i,t}$ the relation between the elasticity and levels. The third line expands $f'(b_{i,t})$ and uses the expressions of the elasticity of demand and the production function. The fourth line simplifies the first term by using the expression of labor supply and the fact that $w_{i,t}l_{i,t} = s_i(b_{i,t})W_t^I L_t$. The fifth line uses the production function to cancel out in the numerator and denominator. The sixth and final line uses the elasticity and levels relationship to recover the effective price pass-through $\rho_{i,t}^{p,\omega}$.

We can combine the above expression with Equation (A31) from Lemma 6 to yield

$$\begin{aligned}
\rho_{i,t}^{p,\omega} &= -\frac{\rho_{i,t}^{p}}{\rho_{i,t}^{w}}\frac{1}{\eta_i(b_{i,t})}\left(\varepsilon_i(a_{i,t})\frac{\partial \ln a_{i,t}}{\partial \ln \omega_i} + 1\right) - \rho_{i,t}^{p} \\
\left(1 + \frac{\rho_{i,t}^{p}}{\rho_{i,t}^{w}}\frac{\varepsilon_i(a_{i,t})}{\eta_i(b_{i,t})}\right)\rho_{i,t}^{p,\omega} &= -\frac{\rho_{i,t}^{p}}{\rho_{i,t}^{w}\eta_i(b_{i,t})} - \rho_{i,t}^{p} \\
\left(\frac{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})}{\rho_{i,t}^{w}\eta_i(b_{i,t})}\right)\rho_{i,t}^{p,\omega} &= -\frac{\rho_{i,t}^{p} + \rho_{i,t}^{p}\rho_{i,t}^{w}\eta_i(b_{i,t})}{\rho_{i,t}^{w}\eta_i(b_{i,t})} \\
\rho_{i,t}^{p,\omega} &= -\frac{\rho_{i,t}^{p} + \rho_{i,t}^{p}\rho_{i,t}^{w}\eta_i(b_{i,t})}{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})}.
\end{aligned}$$

We can substitute this into Equation (A31) that is rearranged to isolate $\rho_{i,t}^{w,\omega}$ to yield

$$\begin{aligned}
\rho_{i,t}^{w,\omega} &= \frac{\rho_{i,t}^{w}}{\rho_{i,t}^{p}}\rho_{i,t}^{p,\omega} + \rho_{i,t}^{w} \\
&= \frac{\rho_{i,t}^{w}}{\rho_{i,t}^{p}}\left(-\frac{\rho_{i,t}^{p} + \rho_{i,t}^{p}\rho_{i,t}^{w}\eta_i(b_{i,t})}{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})}\right) + \rho_{i,t}^{w} \\
&= \rho_{i,t}^{w}\left(-\frac{\rho_{i,t}^{w}\eta_{i,t}}{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})} - \frac{1}{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})}\right) + \rho_{i,t}^{w} \\
&= \frac{-\rho_{i,t}^{w}\rho_{i,t}^{w}\eta_i(b_{i,t}) - \rho_{i,t}^{w} + \rho_{i,t}^{w}\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\rho_{i,t}^{w}\eta_i(b_{i,t})}{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})} \\
&= \frac{\rho_{i,t}^{w}\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) - \rho_{i,t}^{w}}{\rho_{i,t}^{p}\varepsilon_i(a_{i,t}) + \rho_{i,t}^{w}\eta_i(b_{i,t})}.
\end{aligned}$$

Thus, we have proved Proposition 4. □

F Computational Method

We discuss the computational solution method to solve and estimate the model in greater detail. The remainder of the section is organized as follows: Appendix F.1 discusses how $r(\cdot)$ and $s(\cdot)$ are identified, Appendix F.2 outlines how the remainder of the model is solved. We drop the time subscript t in this section for ease of notation.

F.1 Identification of Revenue and Wage Bill Share Functions

Table A14 summarizes the moments used to estimate $r_i(\cdot)$ and $s_i(\cdot)$. These moments only use the data in 1977. To construct these moments in Table A14, we first sort the sample in 1977 by firm revenue, from smallest to largest. For this sorted sample, we calculate each firm's revenue share and evaluate values from the 5th percentile to the 95th percentile in 10-percentile increments, as well as the 97.5th percentile. We take a sampling of the full empirical joint distribution due to U.S. Census disclosure restrictions. Ideally, $r_i(\cdot)$ and $s_i(\cdot)$ are estimated using the full empirical joint distribution. We show the cumulative revenue share for each percentile in Column (2) of Table A14. For each revenue share percentile, we also compute the corresponding cumulative wage bill share, price markup, and wage markdown. To obtain the associated price markup and wage markdowns values, we take local averages within a 5-percentile neighborhood (2.5 percentiles above and below) for all percentiles up to the 85th percentile. For the 95th and 97.5th percentiles, we use a 2.5-percentile neighborhood. We use this procedure to minimize noise.

The information in Table A14 allows us to recover the cumulative revenue and wage bill share

Table A14: Distributional Moments (By Revenue Share)

Perc. (1)	Cum. Revenue Share (2)	Cum. Wage Bill Share (3)	Price Markup (4)	Wage Markdown (5)
5.0	3.319e-04	5.225e-04	1.068	0.501
15.0	2.258e-03	3.286e-03	1.076	0.674
25.0	6.155e-03	8.488e-03	1.075	0.693
35.0	1.302e-02	1.674e-02	1.031	0.816
45.0	2.397e-02	2.916e-02	1.074	0.822
55.0	4.120e-02	5.024e-02	1.042	0.929
65.0	6.734e-02	8.042e-02	1.050	0.960
75.0	1.130e-01	1.314e-01	1.000	1.152
85.0	2.070e-01	2.365e-01	1.033	1.165
95.0	4.579e-01	5.030e-01	1.014	1.144
97.5	5.937e-01	6.586e-01	0.995	1.093

Notes: This table presents the distributional moments used to calibrate the model, based on data from 1977. Column (1) lists percentiles. Columns (2) shows the cumulative revenue share for each percentile. Columns (3) reports the corresponding cumulative wage bill share. Columns (4) and (5) display the associated price markups and wage markdowns. To calculate Columns (4) and (5), we calculate the average markup/markdown within a 5.0-percentile neighborhood (2.5 percentiles above and below) around each revenue share percentile. For the 95th and 97.5th percentiles, we use a narrower 2.5-percentile neighborhood. All figures are rounded in accordance with U.S. Census disclosure requirements.

functions

$$\Lambda_r(i) = \int_{z=0}^{i} r(z)\,dz,$$
$$\Lambda_s(i) = \int_{z=0}^{i} s(z)\,dz.$$

For now, we define $r(\cdot)$ and $s(\cdot)$ by their firm index $i \in [0,1]$. Since we sort the firms by their revenue the index i is their rank with 0 being the lowest and 1 being the highest. Furthermore, since both revenue and wage bill shares are monotone, we also have $s(\cdot)$ in the correct order. With the cumulative distribution functions we can recover $r(\cdot)$ and $s(\cdot)$ through numerical differentiation.

We also possess information on how markups and markdowns vary with i from Table A14. However, we must make an adjustment for two reasons. First, as we discuss in Section 3, the level of the estimate of the price markup is biased (Bond et al., 2021; De Ridder, Grassi and Morzenti, 2025). Second, our model is unable to rationalize firms having either a price markup or wage markdown that is weakly less than 1. Thus, we must shift the distributions of price markups and wage markdowns. We divide the distribution of wage markdowns by the smallest value and then we multiply the distribution a small factor $1 + \zeta_\nu$ where $\zeta_\nu > 0$ and small. We do a similar procedure for price markups but the shift for markups $\zeta_\mu > 0$ is an estimated parameter that is largely determined by the labor share given the wage markdowns. Since we can observe labor shares in the data, given Equation (31) and the wage markdown, there exists a unique price markup that is consistent. We can define the markup and markdown as functions of i.

After the adjustments are made, we fit functions through the samples to get $r(i)$, $s(i)$, μ_i, and ν_i that satisfy Assumptions 6 to 10. Now, we use the results from Proposition 4 to recover $r(\cdot)$ and $s(\cdot)$ as functions of a_i and b_i, respectively. This also leads to the proof of Proposition 5. In the proof we denote A_i as the productivity measure that accounts for the quality and taste shifters.

Proof of Proposition 5. First, we differentiate the log of the markup and markdown functions with respect to i and take the chain rule as follows

$$\frac{\partial \ln \mu_i}{\partial i} = \frac{\partial \ln \mu_i}{\partial \ln a_i}\frac{\partial \ln a_i}{\partial \ln A_i}\frac{\partial \ln A_i}{\partial i},$$
$$\frac{\partial \ln \nu_i}{\partial i} = \frac{\partial \ln \nu_i}{\partial \ln b_i}\frac{\partial \ln b_i}{\partial \ln A_i}\frac{\partial \ln A_i}{\partial i}.$$

We substitute in the expressions from Lemma 5 and Proposition 4 to obtain

$$\frac{\partial \ln \mu_i}{\partial i} = \left(1 - \frac{1}{\rho_i^p}\right)\rho_i^{p,\omega}\frac{\partial \ln A_i}{\partial i}, \tag{A32}$$
$$\frac{\partial \ln \nu_i}{\partial i} = \left(\frac{1}{\rho_i^w} - 1\right)\rho_i^{w,\omega}\frac{\partial \ln A_i}{\partial i}. \tag{A33}$$

In the two expressions above, we can compute the quantities from the left-hand side from the data as well as the elasticities ε_i and η_i since we know markups and markdowns, however, we cannot yet compute the pass-throughs ρ_i^p and ρ_i^w since these depend on the derivatives of markups and markdowns with respect to a_i and b_i, respectively. This also means we cannot directly compute the measured/effective pass-throughs. However, we have enough information to back out the pass-throughs. If we are able to compute the pass-throughs then we can recover $\frac{\partial \ln \omega_i}{\partial i}$ and in turn we can recover $r(\cdot)$ and $s(\cdot)$ as a function of relative prices/wages.

First, we can use the definition of the revenue and wage bill share functions, that is $r(i) = \frac{p_i y_i}{P^I Y}$ and $s(i) = \frac{w_i l_i}{W^I L}$. We take the derivative of the log with respect to i which yields

$$\frac{\partial \ln r(i)}{\partial i} = \frac{\partial \ln p_i y_i}{\partial \ln a_i}\frac{\partial \ln a_i}{\partial \ln A_i}\frac{\partial \ln A_i}{\partial i} = (1 - \varepsilon_i)\rho_i^{p,\omega}\frac{\partial \ln A_i}{\partial i}, \tag{A34}$$
$$\frac{\partial \ln s(i)}{\partial i} = \frac{\partial \ln w_i l_i}{\partial \ln b_i}\frac{\partial \ln b_i}{\partial \ln A_i}\frac{\partial \ln A_i}{\partial i} = (1 + \eta_i)\rho_i^{w,\omega}\frac{\partial \ln A_i}{\partial i}. \tag{A35}$$

These follow from the standard algebra and prior definitions. We substitute Equation (A34) into

Equation (A32) and Equation (A35) into Equation (A33) to obtain

$$\frac{\partial \ln \mu_i}{\partial i} = \left(1 - \frac{1}{\rho_i^p}\right)(1-\varepsilon_i)^{-1}\frac{\partial \ln r(i)}{\partial i} = \left(1 - \frac{1}{\rho_i^p}\right)(1-\mu_i)\frac{\partial \ln r(i)}{\partial i},$$
$$\frac{\partial \ln \nu_i}{\partial i} = \left(\frac{1}{\rho_i^w} - 1\right)(1+\eta_i)^{-1}\frac{\partial \ln s(i)}{\partial i} = \left(\frac{1}{\rho_i^w} - 1\right)\left(\frac{1}{\nu_i} - 1\right)\frac{\partial \ln s(i)}{\partial i}.$$

Thus, we have proved Proposition 5. □

Notice that we can compute all the quantities on the left-hand side of the last two lines of the proof using the data, and the same holds for the intermediate steps and the right-hand side, except for the primitive pass-throughs. Therefore, we can recover the primitive pass-throughs using the results of Proposition 5. With the pass-throughs identified, we can compute the measured/effective pass-throughs and thus we can identify $\frac{\partial \ln A_i}{\partial i}$ through Equations (A32) and (A33). We can use numerical integration to obtain A_i. Note that from Equations (A34) and (A35) we get that

$$\frac{\partial \ln a_i}{\partial i} = \rho_i^{p,\omega}\frac{\partial \ln A_i}{\partial i}, \tag{A36}$$
$$\frac{\partial \ln b_i}{\partial i} = \rho_i^{w,\omega}\frac{\partial \ln A_i}{\partial i}. \tag{A37}$$

Since these are now known, we can rearrange Equations (A34) and (A35) to obtain

$$\frac{\partial \ln r(i)}{\partial \ln a_i} = \frac{\partial \ln r(i)}{\partial i}\left[\rho_i^{p,\omega}\frac{\partial \ln A_i}{\partial i}\right]^{-1} = 1 - \varepsilon_i, \tag{A38}$$
$$\frac{\partial \ln s(i)}{\partial \ln b_i} = \frac{\partial \ln s(i)}{\partial i}\left[\rho_i^{w,\omega}\frac{\partial \ln A_i}{\partial i}\right]^{-1} = 1 + \eta_i. \tag{A39}$$

We can use numerical integration on Equations (A36) to (A39) and use the definition of an elasticity to recover $r(\cdot)$ and $s(\cdot)$ as functions of a_i and b_i, respectively. These steps fully recover the revenue and wage-bill share functions as mappings from revenue rankings to functions of a_i and b_i, respectively.

F.2 Solving the Remainder of the Model

With $r(\cdot)$ and $s(\cdot)$ recovered, we can proceed to solve the remainder of the model. The solution method is standard. The computational method involves making an initial guess of the aggregate state and using the first-order conditions and constraints to check and iterate on this guess until a fixed point is achieved.

Before proceeding with the solution, it is helpful to establish the following result. Equations (A27) and (A28), the relationships between the competition and ideal price indices, can be expressed as follows

$$P^I = P \exp\left(\bar{k}'_P - \int_0^1 \int_{a_i}^{\infty} \frac{r(z)}{z}\, dz\, di\right) = P\bar{K}_P,$$
$$W^I = W \exp\left(\bar{k}'_W + \int_0^1 \int_0^{b_i} \frac{s(z)}{z}\, dz\, di\right) = W\bar{K}_W.$$

These equalities follow from standard rules of integration. How we compute $\bar{K}_P$ and $\bar{K}_W$ depends on whichever approach is numerically more stable to compute. Given these equalities, we guess the competition indices and the factors $\bar{K}_P$ and $\bar{K}_W$, which imply a value for the ideal indices.

Given a draw of firm productivity levels ω_i and a guess of the aggregate state $(P, W, \bar{K}_P, \bar{K}_W, Y, L)$, we can pin down the distribution of firm prices p_i and wages w_i. Note that in this economy in equilibrium we have $Y = C$. We can create a mapping between p_i and w_i given aggregates,

$$p_i = h^{-1}\left(\omega_i f(b_{i,t})\right) P_t.$$

We can use a modification of the pricing rule (29) to recover prices and wages, which is given by

$$\frac{p_i}{P} = \mu(p_i/P)\nu(w_i/W)\frac{w_i}{W}\frac{W}{P}\omega_i^{-1}.$$

We now check that given the distribution of prices and wages as well as the aggregate states we have that

$$\int_{i=0}^{i} r(a_i)\, di = 1,$$
$$\int_{i=0}^{i} s(b_i)\, di = 1.$$

This step is where we update P, W and represents the outer loop. When we update the competitive price indices we restrict it so that it satisfies the production function, demand, and supply constraints. We can combine these constraints starting with the production function as follows

$$y_i = \omega_i l_i,$$
$$\frac{r(a_i)P^I Y}{p_i} = \frac{\omega_i s(b_i) W^I L}{w_i},$$
$$\frac{r(a_i)P\bar{K}_P Y}{p_i} = \frac{\omega_i s(b_i) W\bar{K}_W L}{w_i},$$
$$\frac{r(a_i)}{p_i/P} = \frac{\omega_i s(b_i)}{w_i/W}\frac{L}{Y}\frac{\bar{K}_W}{\bar{K}_P}.$$

Notice that only the ratios $Y/L, P/W, \bar{K}_P/\bar{K}_W$ not the levels matter for the guess. Thus, we update P/W to obtain $\int_{i=0}^{1} r(a_i)\, di = \int_{i=0}^{1} s(b_i)\, di = 1$. When P/W is updated, that implies a new distribution of prices and wages. Thus, these steps represent the inner loop of the solution. The procedure iterates until the errors are sufficiently small to represent an approximated fixed point.