Markups, Labor Share, and Wage Dispersion

February 2023

Abstract

Market powers can affect firm behaviors in labor market. This paper argues that the decrease in the overall wage since 2000 is caused by both rising product markup and falling labor markdown. The significant decrease in wages among lower percentiles has led to an increase in wage dispersion and a decrease in the aggregate labor share. I build a heterogeneous firm model to demonstrate that the decline in wages, especially among the lower percentiles, can be attributed mainly to the monopsony power of firms, which explains the rise in wage inequality and the fall in labor share.

Keywords: Heterogeneous firms, variable markups, wage dispersion, labor share

JEL Codes: D2, E2, J3, L1

1 Introduction

What are the reasons for the decline in the aggregate labor share? What are the sources of the rising wage dispersion? Recent research provides different explanations for the decline in the aggregate labor share such as a fall in the relative price of investment goods, technology change, rising product markups, and productivity dispersion.[1] This paper provides another mechanism that explains this fact: increasing monopsony power in the labor market. Many works focus on the firm's performance in product markets and the implications of product market power (Decker *et al.*, 2016; Gutiérrez & Philippon, 2017; Akcigit & Ates, 2019). However, I show that the firm's decisions in the labor market also plays an important role in explaining the decline in the labor share. Moreover, the rising monopsony power of firms provides another explanation for the increase in wage dispersion.

This paper studies the role of firms' market power in both the product market and the labor market as drivers of the decline in the aggregate labor share and the aggregate wage after 1980. Markup is the firm's ability to charge a higher price over its marginal cost. A rise in markup shows a firm's ability to claim more market power in product markets. Similarly, a wage markdown is defined as the firm's ability to pay a lower wage relative to its marginal revenue product of labor; it shows the worker's ability to bargain over their wages. In a competitive labor market, markdowns would be equal to unity, that is, workers get paid their marginal product of labor. A markdown below unity indicates that the worker has less bargaining power, and, therefore, the firm has monopsony power. This paper argues that the rising aggregate product markup and falling aggregate wage markdown over the past three decades both play an important role in the declining aggregate labor share and aggregate wage. Also, the residual wage dispersion among workers with identical characteristics accounts for a substantial proportion of wage inequality. This wage differentials can be rationalized by the effect from firms heterogeneity of markups and markdowns.

I first follow the methodology by De Loecker *et al.* (2020) to show the evolution of markups based on firm-level Compustat data for the U.S. economy. The aggregate markup increased from 18% in 1980 to 33% in 2018. I further consider factor market distortions in this paper. I use the methodology by Hall (1988) to compute the firm-level labor markdown. The aggregate markdown has declined from 0.99 to 0.85, indicating a rise in firms' monopsony power from 1% to 18%. I then show that the aggregate log wage has fallen 7.5%, the aggregate labor share has declined 23%, and the standard deviation of the wage has risen 77% since 1980. Lastly, I document that the higher markup and lower markdown are associated with a lower

[1] See Karabarbounis & Neiman (2013), Elsby *et al.* (2013), Autor *et al.* (2017a), Kehrig & Vincent (2017), De Loecker *et al.* (2020), and Gouin-Bonenfant *et al.* (2018).

labor share. I also show that the 10th percentile of wages has fallen since 1980, especially after 2000; whereas, the 90th percentile of wages increased slightly over time; both the markup and markdown have a larger impact on the low percentiles of wages than on the high percentiles of wages. Specifically, wages on the left tail of the wage distribution pull down the aggregate wage and drive up the wage dispersion.

I then build a model using the Kimball (1995) aggregator, featuring heterogeneous monopolistically competitive firms facing non-CES demand and workers on-the-job search in the labor market. High-productivity firms face lower demand elasticity and charge higher markups than low-productivity firms. When the product market is perfectly competitive, the high-productivity firms also face less elastic labor supply and have larger firm monopsony power relative to the low-productivity firms. Once the product market becomes noncompetitive, firms operate at lower marginal revenue level and decrease the wages they pay. Moreover, low-productivity firms find it more profitable to claim higher firm monopsony power and reduce wages more than high-productivity firms. As a consequence, low-productivity firms contribute the most to the decline in the aggregate wage and the rise in wage dispersion. Both markup and markdown create a wedge between firms' revenue and cost. As a result, both increasing markup and decreasing markdown cause the decline in the labor share.

To assess the model's performance, I calibrate the model to match the aggregate moments related to product market power: the aggregate markup, markup dispersion, and sales dispersion; and two labor market moments: the aggregate markdown and the replacement rate. The model correctly predict the *trend* and the *level* of the aggregate wage and aggregate labor share as well as the *trend* of wage standard deviation. Moreover, the model capture that, within 6-digit NAICS industry, 80% of the decline in the aggregate labor share stems from the within firm component, firms adjust their labor share internally, and the remaining 20% is explained by the between firm component, sales are reallocated to the firms with low labor share. Most studies of the labor share have only investigated the impact from between sectors reallocation effect, which this paper provides another source of the decline in labor share. Finally, the model can also reproduce the sharp decline in the low percentile wage and the stable trend in the high percentile wage.

How does market power affect the labor share and wage dispersion? I answer this question by simulating three counterfactual analyses one by one: decreasing demand elasticity, falling labor market competitiveness, and increasing productivity dispersion. I first find that, following decreasing demand elasticity, the *markup* channel explains 25% of the decline in the labor share. Secondly, the *markdown* channel contributes 35% to this fall after allowing for falling labor market competitiveness. Finally, the *productivity* channel, by increasing productivity dispersion, causes the most of the decrease—up to 40%.

Most part of the rising markup is the result of the decline in demand elasticity parameters. The fall in demand elasticity allows firms to claim more markup and reduce the wage they pay; hence, the higher the markup a firm has, the more it drops the wage. The counterfactual result shows the markup channel causes the wage level and wage dispersion to decline. On the other hand, I show that both workers' job finding rate and unemployment benefits fell over time, and these represent a lower labor market competitiveness. The decline in labor market competitiveness can explain the fall in the aggregate markdown. A reduction in the labor market competitiveness allows a small firm to reduce wages more and increase its firm monopsony power; therefore, the aggregate wage would drop but wage dispersion would rise. I find that the markdown channel contributes to the sharp decline in the aggregate wage and, in particular, to the low percentile wage the most, especially after 2000. As a result, increasing wage dispersion is mainly driven by the sharp drop in the low percentile wage. Finally, increasing productivity dispersion has a positive effect on wages but the effect is almost offset by the effect from the markup. Higher productivity dispersion allows high-productivity firms to increase their size and pay higher wage; therefore, both the aggregate wage and wage dispersion would rise. I find that an increase in productivity dispersion results in a rise in the aggregate wage and high percentile wages as well as an increase in the wage differential due to increasing high percentile wages. All three groups of parameters contribute differently to the decline in wages and the rise in wage dispersion. However, all these parameters would cause a decline in the labor share.

Literature. This paper is related to several strands of literature. First, it is related to recent literature that studies the increasing market power of a few firms. De Loecker *et al.* (2020) show that the sale-weighted average markup has increased substantially since 1980 in the U.S., especially for a small portion of firms.[2] Autor *et al.* (2017a) argue that the rise in profits is an efficient outcome, reflecting the increasing importance of "superstar" firms. However, the cost of markup could be substantially distorted. Edmond *et al.* (2018) find that the aggregate markup accounts for about two-thirds of welfare costs, and misallocation accounts for about one-third of welfare costs. Edmond *et al.* (2015) argue that opening up to international trade strongly increases competition and reduces markup distortions. In contrast to this literature, I focus on the consequences of product markup on the labor market, especially wage dispersion.

Second, my paper is related to an emerging literature on the decline in the labor share. Karabarbounis & Neiman (2013) show that the decline in the relative price of investment goods induces firms to shift away from labor and toward capital which causes a decline in the labor

[2]The phenomenon is widespread in all the developed countries (De Loecker & Eeckhout, 2018). The firms' granular effect can shape the aggregate phenomena (Gabaix, 2011).

share. Alternately, Elsby *et al.* (2013) advocate for the role of offshoring as an important driver of the labor share decline. Autor *et al.* (2017a) and Kehrig & Vincent (2017) focus on reallocation in product market from firms with relatively high measured labor shares to firms with low labor shares. Gouin-Bonenfant *et al.* (2018) build a model to match Canadian data and show that the main driver of the labor share decline is an increase in firm productivity dispersion. Unlike previous work, in my paper the mechanism is reallocation in both in the product and labor markets. These resource reallocations from low-productivity firms to high-productivity firms have a significant effect on labor share distribution and wage dispersion.

My paper is also related to the literature on wage inequality. Antitrust regulators pay little attention to labor market power despite the labor literature on firms' market power in the labor market (Staiger *et al.* (2010); Falch (2010); Ransom & Sims (2010)). This monopsony in the labor market implies that firms have the ability to pay workers less than their marginal product.[3] Berger *et al.* (2019) develop a tractable quantitative general equilibrium oligopsony model of the labor market and show welfare losses from labor market power that range from 2.9 to 8.0% of lifetime consumption. Unlike Berger *et al.* (2019), this paper deals with market competition that firms face and labor market search friction.[4] The model in this paper features heterogeneous firms engaging in monopolistic competition with non-CES demand.[5] On the other hand, workers are paid their marginal product and are indifferent to where they work in a competitive labor market. Job search and recruiting, however, are not frictionless processes that are required in a case of perfect competition.[6]

The paper proceeds as follows. Section 2 presents the motivating facts to guide the model choice. Section 3 presents the model. Section 4 discusses the model performance. Section 5 provides simulation results for the counterfactuals and explains the effects of markups and markdowns on the labor share and wages. Section 6 concludes.

[3] Helpman *et al.* (2017) build the heterogeneous-firm model of trade and inequality and examine employer-employee data for Brazil's aftermath of trade. Coşar *et al.* (2016) use establishment-level data from Colombia to estimate an open economy dynamic model that links trade to job flows and wages.

[4] Coşar *et al.* (2016) connect the product market with the labor market in an open economy framework but with CES demand.

[5] I follow the Edmond *et al.* (2018) setup to build the firm competition in product markets in which more productive firms are larger and face less elastic demand, and, so, they charge higher markups than less productive firms. More importantly, firms can gain markups by increasing firms' size and claiming more market share; therefore, resource reallocates toward more productive and large firms. Due to economies of scale, larger firms have lower costs, and, therefore, they can increase their market share.

[6] Brown & Medoff (1989) suggest employer size as a measure of the job ladder rung for workers to evaluate which firms they want to hire.

2 Motivating Facts

In this section, I start by presenting the empirical evidence regarding the patterns of aggregate markup, markdown, labor share, and wages from publicly traded U.S. firms over the period 1980 to 2018. I establish the key stylized facts about the relationship between market power, the labor share, and wages that guide the theoretical model developed in Section 3.

Data. I use microdata data from Compustat, which contains information on firm-level financial statements in the U.S. covering from 1980 to 2018. Compustat includes every sector, details up to the 6-digit NAICS level, and provides information on firm-level financial statements that include sales, input expenditures, and capital stock information over a substantial period. Markup is defined as price over marginal cost. I follow the methodology in De Loecker & Warzynski (2012) to estimate firm-level markup, and the information of firms in Compustat allows me to measure markups, as in De Loecker *et al.* (2020).[7] On the other hand, markdown is defined as firm wage over its marginal revenue product of labor. I use the methodology in Hall (1988) to obtain industry-wide labor cost shares to measure the firm-level markdown.[8]

To compute the labor share, I follow Keller & Yeaple (2009) and use firms' sales minus staff expenses and operating income to construct material expense. I then compute the value-added labor share as staff expense divided by sales minus material expense.[9] The value-added share used in the calculations below is the denominator of the relevant labor share measure. Thus, I consider the payroll-to-value-added ratio as the firm's sales weight. In the following context, I refer to value-added as firm-level sales to avoid confusion. Lastly, the firm-level wage is computed by the logarithm of the firm-level wage bills divided by the number of employees.[10]

[7] Firms listed in Compustat are publicly-traded firms, and they account for only 29% of private US employment (Davis *et al.*, 2006). De Loecker *et al.* (2020) deals with the selection of the publicly-traded firms in two ways. First, they use the U.S. Censuses to compute markups. The results from the U.S. Censuses are consistent with the results from Compustat. However, except for the manufacturing sector, which contains establishment-level data on sales in addition to very comprehensive data on inputs (the total labor wage bill, capital, materials, and so on), most of the other sector censuses (retail, wholesale, etc.) only contain data on establishment-level sales, the wage bill, and not other nonlabor inputs. Second, they use the population weights of each sector to adjust the weights in the Compustat sample and account for any bias due to the sectoral composition.

[8] Although Compustat does not have comprehensive data for the wage bill, the evolution of markups for those reported companies is consistent with the evolution of markups for the entire sample. It is still worth glancing at the markup's effect on labor share and wages, but with caution. Whether the distribution of the markup would affect wage dispersion is still unclear in the literature. To my knowledge, this is the first paper to try to understand the effect of the markup distribution on wage inequality.

[9] This income approach is the same methodology as in the System of National Accounts and is also used in Gouin-Bonenfant *et al.* (2018).

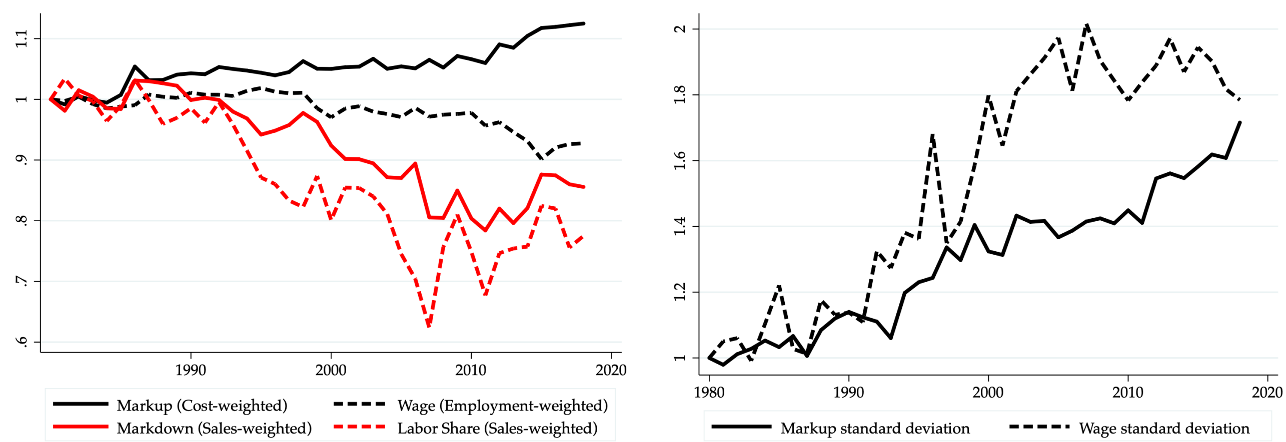
[10] Appendix A documents a more detail estimation procedure.

Empirics. I define the aggregate level of markup, wage, labor share, and markdown using following weighting scheme.

$$\mathcal{M}_t = \sum_i \mathcal{CW}_{it}\mu_{it}, \quad W_t = \sum_i \mathcal{EW}_{it}w_{it}, \quad LS_t = \sum_i \mathcal{SW}_{it}ls_{it}, \quad \mathcal{M}_t^d = \sum_i \mathcal{SW}_{it}\mu_{it}^d$$

where the aggregate markup $\mathcal{M}_t$ is defined as the cost-weighted ($\mathcal{CW}_{it}$) firm-level markup μ_{it} in a given year. Similarly, the aggregate wage W_t is the employment-weighted ($\mathcal{EW}_{it}$) firm-level wage w_{it}, the aggregate labor share LS_t is the sales-weighted ($\mathcal{SW}_{it}$) firm-level labor share ls_{it}, and the aggregate markdown $\mathcal{M}_t^d$ is the sale-weighted firm-level markdown μ_{it}^d.

Figure 1 The Evolution of Markup, Wage, Labor Share, and Markdown



(a) Average

(b) Standard Deviation

Figure 1 panel (a) shows the evolution of the aggregate markup, wage, labor share, and markdown, and panel (b) shows the evolution of the standard deviation of the markup and wage from 1980 to 2018. I normalized all the variables to one in 1980. The aggregate markup has increased from 1.18 in 1980 to 1.33 in 2018[11], which is around a 12.5% increase, and the standard deviation of the markup increased around 70% through this period. The aggregate wage slightly increased from 3.96 log points in 1980 to 4.01 log points in 1998 and then dropped to 3.67 log points in 2018, which is a 29% decline in the wage level. Meanwhile, the standard deviation of the wage has nearly doubled. Lastly, the aggregate labor share fell by 23%[12], and the aggregate labor markdown dropped by 15% from 1980 to 2018. Therefore, as both product

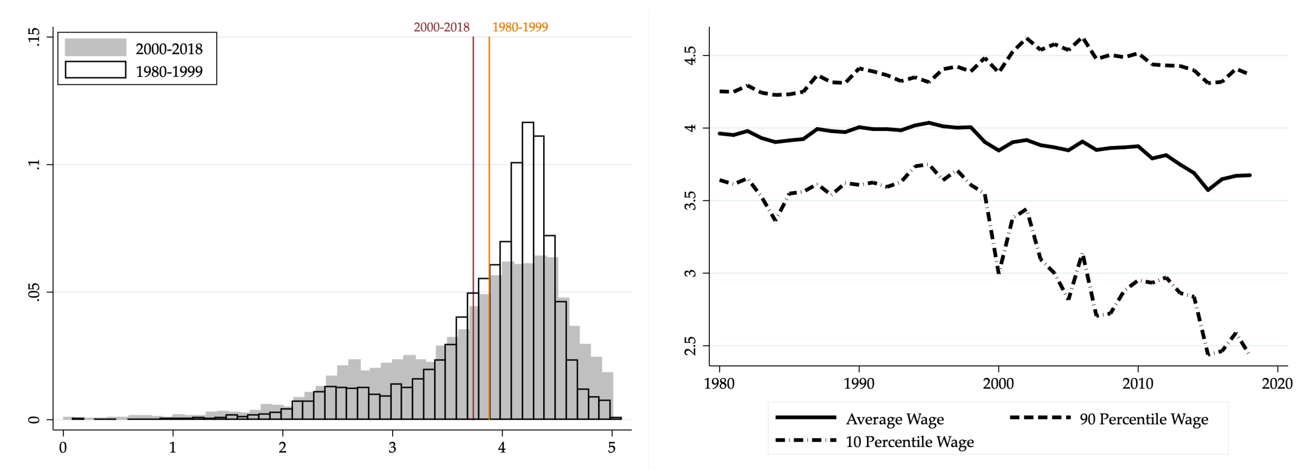
[11] De Loecker *et al.* (2020) conclude that the sales-weighted average markup rose from 21% in 1980 to 61% in 2016. By contrast, I estimate the cost-weighted average markup, which is similar to Edmond *et al.* (2018) and is inline with model implication. For further detail on difference between the sales-weighted and the cost-weighted can be refered to Edmond *et al.* (2018).

[12] Although real weekly earnings from the BEA is a slightly increases over this period and the labor share is a decline by 11%, the presented data are calculated differently from BEA. First, the data is less inclusive than the

market power and labor market power grew over 40 years, the aggregate labor share and the aggregate wage declined over time, and the wage dispersion rose through out the period.

Figure 2 depicts the unweighted wage distribution before and after 2000, and the weighted wage percentile over the time. The decline in the aggregate wage, especially after 2000, came from the low percentile of wages. Specifically, the 10th percentile declined substantially since 2000 from 3.5 log points to 2.5 log points; whereas, the 90th percentile of wages increased slightly between 1980 and 2018. The wage distribution expanded on both the left and right tail of the distribution after 2000. Most importantly, the distribution features a thicker left tail, indicating that most firms pay lower wages than average. Although only a small fraction of firms pay higher wages than before 2000, the unweighted aggregate wage still declined.

Figure 2 Wage Distribution and Employment-weighted Percentile Wage



(a) Unweighted Wage Distribution

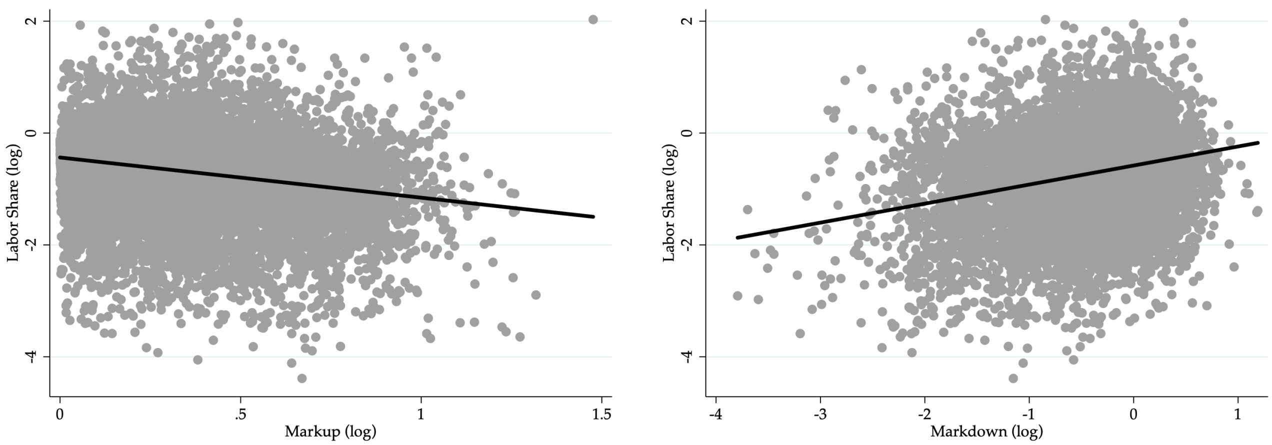
(b) Employment-weighted Percentile Wage

Facts. I now show the empirical evidence that guides the theoretical model developed below. Figure 3 shows the scatter plots of the correlation of the firm-level markup and markdown with its value-added labor share. A higher markup and a lower markdown are associated with a lower labor share. In Table 1, I report the regression coefficients of the log of the labor share on the log of the firm's markup and markdown; all columns are clustering at the firm level. The

BEA measure for labor cost (in the aggregate, the BEA includes self employed or proprietor's labor compensation (corporations, partnerships and sole proprietorships), and some firms in our sample do not include commissions, bonuses or incentive compensation) Second, the calculation uses the weighted aggregate, which reflects the firms' wage payments over time. Finally, even though I impute the material expense to compute the share of labor expenditure in value added, the reported labor share, from 0.56 to 0.43, is a lot lower than the labor share reported in the macro literature. Although there are some discrepancies between the Compustat data and the Census data, which represents more general cases, the patterns of the top firms and the aggregate outcome are similar as documented in De Loecker *et al.* (2020). The result cannot be referred to the general economy but, at least, it presents the parts of economy that significantly affect the aggregate outcome.

first three specifications replicate the results in De Loecker *et al.* (2020), without considering the markdown effect. Failing to control for the markdown effect leads to the coefficients of markup being downward biased. The fourth to sixth columns include the markdown effect. The effects of a markup on the labor share are larger than the first three specifications, which are not controlling the markdown effects. The results consistently indicate a negative coefficient of around -0.85 to -0.9 in markup and a positive coefficient of around 0.37 to 0.4 in markdown. Therefore, as a firm's markup increases by 10%, its labor share decreases by around 9%; whereas, as a firm's markdown decrease by 10%, its labor share decreases by around 4%. It can thus be suggested that both increasing the aggregate markup and decreasing the aggregate markdown could be a reason for the decline in the aggregate labor share.

Figure 3 Effects of Markup and Markdown on Labor Share



(a) Markup (b) Markdown

To assess the effect of market power on wages, I test the prediction that heterogeneous effects of markups and markdowns lead to a decline in the aggregate wage. I estimate the following specifications

$$\log(w_i) = \eta_s + \gamma_t + \log(\mu_i) + \log(\mu_i^d) + \log(l_i) + \log(y_i)$$

where μ_i and μ_i^d denote the firm-level markup and markdown, respectively, w_i is the firm-level log wage, γ_t is year fixed effects, and η_s is industry fixed effects. I also run two separate regressions for firm wages below the 25 percentile and above the 75 percentile. The results are shown in Table 2. The first column reports the effects of markup and markdown on the overall firm-level wage. They are roughly the same; an increase of 10% in markup or a decrease of 10% in markdown results in a 8.5% decline in wage. Most importantly, these effects on the wage are mostly driven by the lower-wage percentile. As we move up to the higher percentile, both

Table 1 The effects of firm-level markups and markdowns on firm-level labor shares

	Labor Share (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Markup (log)	-0.719***	-0.656***	-0.570***	-0.877***	-0.847***	-0.906***
	(0.0647)	(0.0650)	(0.0751)	(0.0602)	(0.0609)	(0.0692)
Markdown (log)				0.399***	0.387***	0.375***
				(0.0247)	(0.0246)	(0.0207)
Year F.E.		X	X		X	X
Industry F.E.			X			X
Observations	23282	23282	23282	23282	23282	23282
R^2	0.053	0.070	0.513	0.148	0.155	0.570

Notes: Columns (1) and (4) start the estimation without controlling for year fixed effects and industry fixed effects. Columns (2) and (5) add control for year fixed effects. Columns (3) and (6) add control for both year fixed effects and industry fixed effects. Standard errors are clustered at the firm level for all regressions in parentheses. Levels of significance: *10%, **5%, ***1%.

Table 2 The effects of firm-level markups and markdowns on firm-level wages

	(1)	(2)	(3)
	W	$< W_{25}$	$> W_{75}$
Markup (log)	-0.855***	-0.885***	-0.456***
	(0.0402)	(0.0665)	(0.0517)
Markdown (log)	0.875***	0.828***	0.505***
	(0.0122)	(0.0191)	(0.0304)
Employment (log)	-0.871***	-0.786***	-0.551***
	(0.0116)	(0.0184)	(0.0293)
Sales (log)	0.854***	0.772***	0.525***
	(0.0112)	(0.0176)	(0.0287)
Year F.E.	X	X	X
Industry F.E.	X	X	X
Observations	23368	5858	5772
R^2	0.871	0.867	0.825

Notes: Standard errors are clustered at the firm level for all regressions in parentheses. Levels of significance: *10%, **5%, ***1%.

markup and markdown effects are half of the general effects. According to these data, I can infer that the effect of markup and markdown on the decline in the aggregate wage mainly comes through the channel of a decline in the low-percentile wage.

3 Model

This section lays out the model to connect the firms' product market decisions with its labor market decisions. Heterogeneous monopolistically competitive firms face non-CES demand as described by Edmond *et al.* (2018). Firms hire workers who can search while employed or unemployed in the spirit of Burdett & Mortensen (1998). The model serves two purposes. First, it allows me to see the effect of different market powers on wages and the labor share. Second, it provides intuition for how a firm's markup alters a firm's behavior in the labor market.

The Economy. In the model, the economy consists of a measure L of homogeneous, infinitely lived worker-consumers with preferences over final consumption. The final good is produced by perfectly competitive firms using a bundle of differentiated intermediate inputs. The differentiated intermediate inputs are produced by monopolistically competitive firms using labor. There is no aggregate uncertainty. I focus on characterizing the steady-state equilibrium.

Preferences. Each period t, agents derive utility from the consumption of a composite good, C_t and they do not save. The representative worker-consumer maximizes the expected present value of their utility stream

$$\int_0^\infty e^{-rt} C_t \tag{1}$$

subject to the budget constraint

$$C_t = W_t L_t + \Pi_t$$

where W_t denotes the aggregate real wage, L_t denotes the aggregate labor supply, and Π_t denotes the aggregate firm profits, net of the cost of creating new firms. Because firms are owned by the representative worker-consumer they use the one-period discount factor $\beta = \frac{1}{1+r}$ to discount future profit flows where r is real interest rate.

Labor Market. Assume each worker-consumer consumes either wage w_t when employed or unemployment benefits b_t when unemployed. The homogeneous worker-consumers are receiving wage offers at the Poisson rate λ_t^u when unemployed and with the rate λ_t^e when employed. Wage offers are random draws from a cumulative wage offer distribution $F(w)$ with upper support w_{max}, in which wages are increasing with firm productivity. Workers can accept or reject a wage offer. Worker-consumers discount the future at a real interest rate r and follow a reservation wage strategy where the minimum accepted wage is denoted w_t^*. The value of being employed J with the current wage w_t is:

$$rJ(w_t) = w_t + \lambda_t^e \int_{w_t}^{w_{max}} [J(z) - J(w_t)]\, dF(z) - \delta_t(J(w_t) - U) \tag{2}$$

where U is the value of unemployment. The first term is the current wage, the second term represents the possibility of receiving more lucrative on-the-job offers with probability λ_t^e, and the third term is the value of moving into unemployment with probability δ_t after exogenous job (firm) destruction.

The value of unemployment is

$$rU = b_t + \lambda_t^u \int_{w_t^*}^{w_{max}} [E(z) - U]\, dF(z) \tag{3}$$

An unemployed worker-consumer receives benefits b and job offers at a rate λ_t^u. The reservation wage is characterized by:[13]

$$w_t^* = b_t + (\lambda_t^u - \lambda_t^e) \int_{w_t^*}^{w_{max}} \frac{1 - F(z)}{r + \delta_t + \lambda_t^e[1 - F(w_t)]} dz, \tag{4}$$

which is the sum of the flow benefits in unemployment and the value of continued search in unemployment. Focusing on the effects of firms' market power on wage decisions, I assume that the job arrival rates are the same when workers are unemployed or employed, i.e, $\lambda_t^u = \lambda_t^e = \lambda_t$[14], which implies $w_t^* = b_t$.

In the steady state, the unemployment rate u can be derived from the condition that the outflow from unemployment equals the inflow into unemployment

$$\lambda[1 - F(w^*)]u = \delta(L - u)$$

Rearranging yields

$$u = \frac{\delta}{\delta + \lambda[1 - F(w^*)]} \tag{5}$$

[13] See Appendix B for the derivation.

[14] In a job ladder model, the latter is decreasing in the difference $\lambda^u - \lambda^e$, because worker-consumers are giving up less in terms of search efficiency when moving out of unemployment. In other words, worker-consumers would be willing to accept lower-paying jobs since they can climb up the job ladder without a throwback in unemployment. Similarly, r and w decrease the value of additional search because workers become impatient, and high wage offers have a lower duration, respectively. To abstract the employment and/or wage decisions from worker's heterogeneity, in this context, unemployment is not a search for a good match, but rather bad luck. Consequently, workers accept the first wage offer above their unemployment compensation.

The stationary distribution of employment over wages implies that we can pin down the realized distribution of wage $G(w)$ from the following equation:

$$(L-u)G(w)[\delta+\lambda[1-F(w)]] = u\lambda[F(w)-F(w^*)]$$

The left-hand side of the equation represents the flow of workers with wages below w, who either flow to unemployment or receive a better job offer. The right-hand side of the equation represents the flow of workers moving into the job with wage w. Rearranging yields[15]

$$G(w) = \frac{F(w)-F(w^*)}{1-F(w^*)}\frac{\delta}{\delta+\lambda\left[1-F(w)\right]} \tag{6}$$

The wage distribution is affected by exogenous job (firm) destruction that moves worker-consumers into unemployment, where they subsequently accept any offer above their reservation wage w^*, as well as by the number of job offers. The labor supply denotes the measure of workers per firm earning a wage w and can be written as

$$l(w|w^*, F(w)) = \lim_{\epsilon\to\infty}\frac{G(w)-G(w-\epsilon)}{F(w)-F(w-\epsilon)}(L-u)$$

I define $F(w) = F(w^-)+v(w)$ where $v(w)$ is the mass of firms offering w and w^- is the wage below but not including w. Denoting $k=\lambda/\delta$, I can rewrite the labor supply to firms that offer a wage w as

$$l(w|w^*, F(w)) = \frac{Lk}{\left[1+k[1-F(w)]\right]/\left[1+k[1-F(w^-)]\right]} \quad \text{if } w \geq w^* \tag{7}$$

otherwise $l(w|w^*, F(w)) = 0$ if $w < w^*$. The expression in equation (7) can be written as

$$l(w|w^*, F(w)) = Lk\left[1+k[1-F(w)]\right]^{-2} \tag{8}$$

which is increasing in the wage offer distribution. Firms that offer higher wages can attract more workers and an increase in the wages offered by all other firms decrease the labor supply to that firm. The elasticity of labor supply $\eta(w)$ to the firm and its labor monopsony power $\nu(w)$ can be written as

$$\eta(w) = \frac{d\log l}{d\log w} = \frac{2kF'(w)}{[1+k(1-F(w))]}\times w \tag{9}$$

[15]See Appendix B for the derivation.

$$\nu(w) = \frac{1+\eta(w)}{\eta(w)} \tag{10}$$

$\nu(w)$ reflects firms' monopsony power in labor market and above unity. Define labor mark-down $\mu^d(w)$ as the inverse of $v(w)$, $\mu^d(w) = \frac{1}{v(w)}$, this reflects the extent to which wages are marked down.[16] As long as the wage offer distribution is decreasing in wage, the firms offering higher wage would face less elastic labor supply and extract higher market power in the labor market and mark down more on wage.

Final Good Producers. Let Y_t denote the aggregate production of the final good. This can only be used for consumption C_t, which implies

$$Y_t = C_t$$

The final good Y_t is produced by perfectly competitive firms using a bundle of differentiated intermediate inputs $y_t(n)$ for $n \in [0, N_t]$, where N_t denotes the mass of available varieties. This bundle of inputs is assembled into final goods using the Kimball aggregator, i.e.

$$\int_0^{N_t} \Upsilon\left(\frac{y_t(n)}{Y_t}\right) dn = 1 \tag{11}$$

Normalizing the price of the final good to 1, final good producers choose $y_t(n)$ given $p_t(n)$ to maximize profits subject to technology (11).

$$Y_t - \int_0^{N_t} p_t(n) y_t(n) dn$$

The optimality condition for this problem gives rise to the demand curve facing each intermediate producer

$$p_t(n) = \Upsilon'\left(\frac{y_t(n)}{Y_t}\right) D_t \tag{12}$$

where

$$D_t = \left(\int_0^{N_t} \Upsilon'\left(\frac{y_t(n)}{Y_t}\right) \frac{y_t(n)}{Y_t} dn\right)^{-1} \tag{13}$$

is a demand index. Denote $q_t = y_t(n)/Y_t$ as relative output.

One set of preferences that has been used often in the variable markups literature is the one

[16] In wage-bargaining models, the wages in a match between a worker with reservation wage w^* and a firm with productivity z, and $1/v(w)$ is the bargaining power of the worker.

first studied by Kimball (1995).[17] Following the Klenow & Willis (2016) specification

$$\Upsilon(q) = 1 + (\sigma - 1)\exp\left(\frac{1}{\varepsilon}\right)\varepsilon^{\frac{\sigma}{\varepsilon}-1}\left[\Gamma\left(\frac{\sigma}{\varepsilon}, \frac{1}{\varepsilon}\right) - \Gamma\left(\frac{\sigma}{\varepsilon}, \frac{q^{\varepsilon/\sigma}}{\varepsilon}\right)\right] \tag{14}$$

with $\sigma > 1$ and $\varepsilon \geq 0$ and where $\Gamma(s, x)$ denotes the upper incomplete Gamma function:

$$\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$$

The Klenow-Willis specification in (14) gives

$$\Upsilon'(q) = \frac{\sigma - 1}{\sigma}\exp\left(\frac{1 - q^{\frac{\varepsilon}{\sigma}}}{\varepsilon}\right) \tag{15}$$

which implies the demand elasticity

$$\sigma(q) = -\frac{\Upsilon'(q)}{\Upsilon''(q)} = \sigma q^{-\frac{\varepsilon}{\sigma}} \tag{16}$$

which in turn implies the markup function

$$\mu(q) = \frac{\sigma(q)}{\sigma(q) - 1} \tag{17}$$

When $\varepsilon = 0$, this reduces to the familiar CES constant markup $\mu = \sigma/(\sigma - 1)$. When $\varepsilon > 0$, larger firms find it optimal to choose higher markups. A firm's markup increases with its relative size and is determined by ε/σ, which is critical in determining how markups vary with productivity and competition.

Intermediate Input Producers. Each variety n is produced by a single firm. Firms are created by paying a sunk cost κ in units of labor. On entry, a new firm obtains a one-time productivity draw $z \sim P(z)$, which is independently distributed and drawn from a Pareto distribution $P(z) = 1 - (z_{min}/z)^{\xi}$ for $z \geq z_{min} > 0$ and $\xi > 1$.[18] Firms exit with exogenous probability δ_t each period. I focus on a symmetric equilibrium where producers with the same z will make the same decisions, so henceforth I will index firms by z. A firm with productivity z hires labor

[17] The CES case is a special case, where $\Upsilon(q_t) = q_t^{\frac{\sigma-1}{\sigma}}$, the demand index is a constant $D_t = \sigma/(\sigma - 1)$ and equation (12) reduces to the familiar constant elasticity demand curve $p_t(n) = (y_t(n)/Y_t)^{\frac{-1}{\sigma}}$.

[18] The Pareto distribution is not only tractable but, together with our other assumptions, implies a Pareto firm-size distribution, which provides a reasonable approximation to observed data (Axtell (2001)).

l_t to produce output according to

$$y_t(z) = zl_t^{\alpha} \tag{18}$$

where $\alpha < 1$, that is firms are subject to decreasing returns to scale.

Firms' Problem. In the steady state, each firm maximizes profits by taking as given the production function (18), the demand curve (12), and the labor supply (8). I will focus on size-invariant equilibria in which wages are increasing in firm productivity as in Coles & Mortensen (2016) and Gouin-Bonenfant *et al.* (2018).[19] I can write the static profits of a firm of type z as

$$\pi(z) = \max_{y \geq 0} \left[\Upsilon' \left(\frac{y(z)}{Y} \right) Dy(z) - w(z) \left(\frac{y(z)}{z} \right)^{\frac{1}{\alpha}} \right] \tag{19}$$

Let $y(z)$ denote the solution to the firm's static problem and $q(z)$ is the solution for its relative output. The firm's price $p(z)$ can be written as a markup $\mu(q(z))$ over marginal cost

$$p(z) = \mu(q(z)) \times mc(z) \tag{20}$$

where

$$mc(z) = \nu(z) \times \frac{1}{\alpha} \left(\frac{y(z)}{z} \right)^{\frac{1}{\alpha}} \frac{1}{y(z)} \times w(z)$$

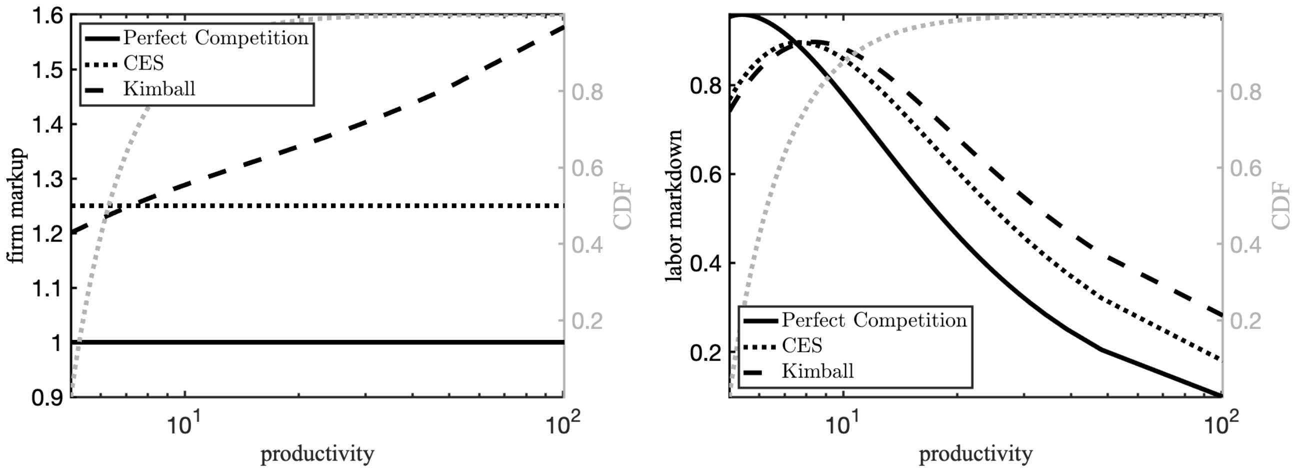
Since the firm has market power in the labor market, its marginal cost of labor accounts for both the wage and the additional cost associated with increasing wages. Rewriting (20), the firm's optimality condition is: marginal revenue product of labor equals marginal expense of labor

$$\underbrace{\underbrace{\frac{p(z)}{\mu(z)}}_{\text{MR}} \times \alpha \underbrace{\frac{y(z)}{l(z)}}_{\text{MPL}}}_{\text{MRPL}} = \underbrace{\nu(z) \times w(z)}_{\text{ME}} \Rightarrow \mu^d(z) = \frac{1}{\nu(z)} = \frac{w(z)}{MRPL}$$

Markups and Markdowns. Figure 4 shows the firm markup and labor markdown for different sizes of firms in perfect competition, monopolistic competition with CES, and monopolistic competition with Kimball demand. The gray dash line represents the productivity distribution, which is the Pareto distribution. When the product market is perfectly competitive, all firms operate at the same marginal revenue level. Therefore, the gap between the wage and the

[19]The focus on size-invariant equilibria means that a firm with productivity z and size l offers a wage $w(z)$ that does not depend on its size l but depends on its productivity z. The focus on equilibria in which wages are increasing in firm productivity means that the equilibrium policy function of the firm satisfies $w(z') > w(z)$ whenever $z' > z$.

Figure 4 Product Market Competitiveness



(a) Firm Markup

(b) Labor Markdown

marginal product of labor reflects the labor markdown. As mentioned earlier, the firms offering higher wage would face less elastic labor supply and markdown more on the wages they pay. Therefore, the most productive firm depresses the wage more than the least productive firm even if the most productive firm pays the highest wage. This is the Burdett & Mortensen (1998) mechanism that the elasticity of the labor supply to a firm decreases as the wage increases.

However, if the product market is monopolistically competitive, all firms are allowed to charge markups. Allowing firms to extract markup would uniformly lower the wages they pay relative to the competitive market case. Nevertheless, the effects of markups on labor markdowns are subtle but nontrivial. Given that all firms' markups move from a unity to larger than a unity, markups have a differential impact on labor markdowns for each firms. Low-productivity firms further extract their markdowns more relative to the competitive market case. On the other hand, High-productivity firms, instead, extract less labor markdowns in the monopolistic competition case. Figure 4 (b) shows that low-productivity firms increase their monopsony power and decrease labor markdown once the product market become noncompetitive. The gray dash line indicates the cumulative distribution of productivity, and it shows that the majority of firms are at the lower range of productivity. Therefore, a large fraction of firms extract larger monopsony power in the labor market once they have market power in the product market.

Labor Share and Wage. Using (20), the firm-level labor share can be written as

$$\frac{w(z)l(z)}{p(z)y(z)} = \frac{\alpha}{\theta(z)} \tag{21}$$

where $\theta(z) = \mu(z)/\mu^d(z)$ is the firm-level wedge. In a perfectly competitive product market and labor market, the firm-level wedge is unity. However, imperfect competition creates a wedge between labor share and output elasticity. The wedge term acts like a tax that reduces the output level compared to the competitive market level. Both increasing markup and monopsony power would drive up the wedge and reduce the labor share.

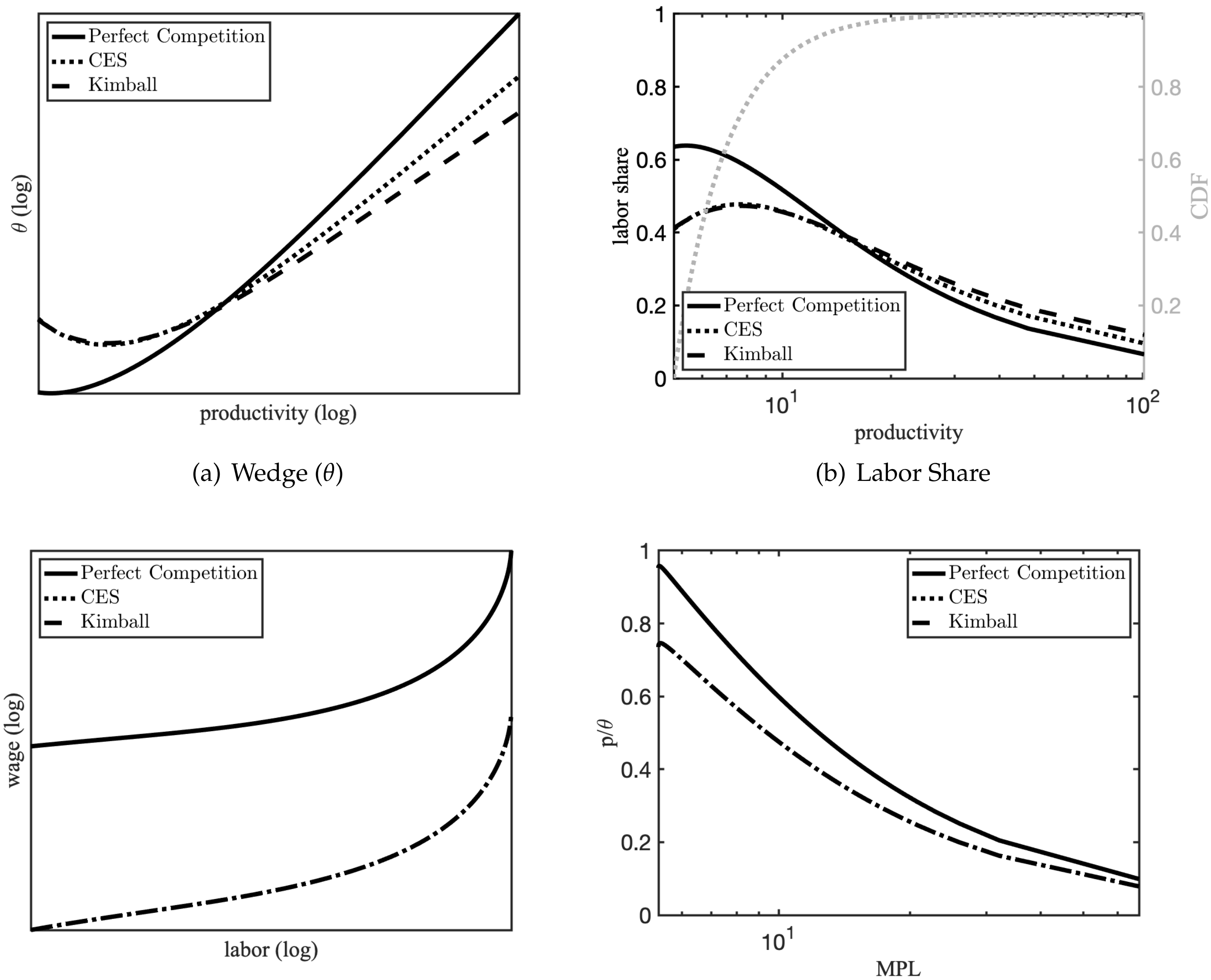
As mentioned earlier, firms with different sizes have a different strategies to create a wedge that extracts profit. Similarly, given that firms face the same labor supply and operate at the same output level in the different competitive markets, the wage that the firm pays is related to its price and wedge, $w \sim p/\theta$. Figure 5 shows the wedge term, labor share, and wage-to-MPL ratio in different models for the product market. Including product market power allows small firms to extract more product market power and labor market power; whereas, large firms increase their product market power at the expense of lower labor market power. The majority of firms are small, and this will translate to a large fraction of low-productivity firms lowering their labor share, but only a small fraction of high-productivity firms increasing their labor share.

The small firms claim higher monopsony power by reducing their wage-to-MPL ratio. Under perfect competition, the least productive firm pays at its MPL. However, the least productive firm only pays 80% of its MPL once the product market becomes noncompetitive. The ratio decreases when productivity increases, but the difference between competitive and noncompetitive markets also reduces by the firm size. This indicates that low-productivity firms play a vital role in wages when the product market is noncompetitive. In monopolistic competition, the wage that every firm pays drops, but wages in small firms drop more than wages in large firms. There is no difference between homogeneous and heterogeneous markup in wages because Kimball demand allows firms to claim higher markup by moving down along the demand curve; this will alters firms MRPL and its wage offer distribution. Therefore, only the labor supply elasticity would change at a given level of wage and labor supply.[20]

Free Entry. Let M_t denote the mass of entrants in period t. Free entry drives the expected profits of potential entrants to zero. The sunk entry cost κW_t is paid prior to the realization of the productivity draw z. The free entry condition determining the equilibrium number of

[20]One way to break the link is by introducing a demand shifter in Kimball demand. The main focus of this paper is to understand the effect of market competition on the labor share and the wage; therefore, adding a demand shifter allows the market to reallocate resources differently, which complicates the analysis. For quantitative reasoning, using Kimball demand can produce a reasonable markdown and labor share since both CES and perfect competition require a large aggregate markdown and a low labor share to match the data.

Figure 5 Labor Share and Wage



(a) Wedge (θ)

(b) Labor Share

(c) Wage

(d) Wage-to-MPL ratio

active firms is given by:

$$\kappa W_t = \int \left(\beta \int_{i=1}^{\infty} (\beta(1-\delta_t))^{i-1} \pi(z)_{t+i} di \right) dP(z) \tag{22}$$

where the aggregate real wage W_t is an employment-weighted arithmetic average of firm-level wage

$$W_t = \int w_t(z) \frac{l_t(z)}{\tilde{L}} dn(z) \tag{23}$$

The expression in equation (22) can be written as

$$\kappa W_t = \int \left(\beta \int_{i=1}^{\infty} (\beta(1-\delta_t))^{i-1} \left(1 - \frac{\mu^d_{t+i}}{\mu_{t+i}} \right) p_{t+i}(z) y_{t+i}(z) di \right) dP(z) \tag{24}$$

A firm's incentives to enter are determined by its operating profits; therefore, it is a function of markups, markdowns, and the firm's overall sales.

Aggregation. Let Z_t denote the aggregate productivity of this economy, implicitly defined by an aggregate production function that relates the total amount of final goods Y_t to the total amount of labor $\tilde{L}_t$ used in production:

$$Y_t = Z_t \tilde{L}_t^{\alpha} \tag{25}$$

where the aggregate productivity Z_t can be expressed in terms of firm-level productivities z according to

$$Z_t = \left(\int \left(\frac{q_t(z)}{z} \right)^{\frac{1}{\alpha}} dn(z) \right)^{-\alpha} \tag{26}$$

Let $\mathcal{M}$ denote the aggregate markup of this economy, $\mathcal{M}^d$ denote the aggregate markdown of this economy, and Θ denote the aggregate wedge; therefore, the aggregate labor share is implicitly defined by

$$\frac{W_t \tilde{L}_t}{Y_t} = \alpha \frac{\mathcal{M}^d_t}{\mathcal{M}_t} = \frac{\alpha}{\Theta_t} \tag{27}$$

The aggregate markup and markdown prevent firms from reaching full capacity and act like a wedge between the labor share and labor productivity. The aggregate wedge Θ_t is a cost-weighted arithmetic average of firm-level wedge $\theta_t = \mu_t / \mu^d_t$

$$\Theta_t = \int \theta_t(z) \frac{w_t(z) l_t(z)}{W_t \tilde{L}_t} dn(z) \tag{28}$$

In turn, the aggregate markup is a cost-weighted arithmetic average of firm-level markups

$$\mathcal{M}_t = \int \mu_t(z) \frac{w_t(z) l_t(z)}{W_t \tilde{L}_t} dn(z) \tag{29}$$

and the aggregate markdown is a sales-weighted arithmetic average of firm-level markdowns

$$\mathcal{M}^d_t = \int \mu^d_t(z) \frac{p_t(z) y_t(z)}{Y} dn(z) \tag{30}$$

The aggregate labor share in (27) can be rewritten as a sales-weighted arithmetic average of the firm-level labor share

$$LS_t \equiv \frac{W_t \tilde{L}_t}{Y_t} = \int \frac{w(z)l(z)}{p(z)y(z)} \frac{p_t(z)y_t(z)}{Y} dn(z) \tag{31}$$

Equilibrium. Let $N_t = \int dn_t(z)$ denote the overall mass of firms. An equilibrium is a sequence of firm prices $p_t(z)$ and wage $w_t(z)$; and allocations $y_t(z)$ and $l_t(z)$; as well as a mass of new entrants M_t, aggregate real wage W_t, aggregate output Y_t, consumption C_t, and labor supply L_t, and a measure of firms N_t such that firms and consumers optimize and the labor and goods markets are all clear.[21]

The total mass of firms evolve according to

$$\dot{N} = M_t - \delta_t N_t \tag{32}$$

Labor market clearing requires

$$L_t = \int l(w_t) dn(z) + \kappa M_t \tag{33}$$

Similarly, goods market clearing requires

$$Y_t = C_t \tag{34}$$

4 Quantitative Analysis

This section outlines the calibration strategy, the model performance, and its implications for the labor share and wage. I start by discussing the calibration strategy. I then compare the model results with empirical evidence and discuss the labor share and wage dispersion.

Calibration Strategy. I calibrate the model to match the aggregate markup, the aggregate markdown, and the standard deviation of markup, which are represented by the solid line in Figure 1. I leave the aggregate wage, the aggregate labor share, and the standard deviation of the wage as untargeted moments, which are shown by the dash line. Therefore, I can understand the effect of both product market power and labor market power on the decline in the aggregate labor share, decline in the aggregate wage, and increase in wage dispersion.

[21] See Appendix B for a detailed solution

Table 3 Parameters

Parameter	Description			Source		
Assigned						
δ	Job destruction rate	Longitudinal Business Database				
ρ	$\frac{\textit{Average Weekly Benefit}}{\textit{Average Weekly Wage}}$	Department of Labor				
Joint Calibration		1980	1990	2000	2010	2018
ξ	Pareto tail	6.12	5.11	4.54	4.15	4.04
σ	Demand elasticity	8.98	7.11	6.65	5.72	4.82
ε/σ	Superelasticity	0.45	0.34	0.31	0.27	0.24
λ	Job arrival rate	0.15	0.16	0.14	0.06	0.07
b	Unemployment benefit	2.99	2.98	2.87	2.67	2.57

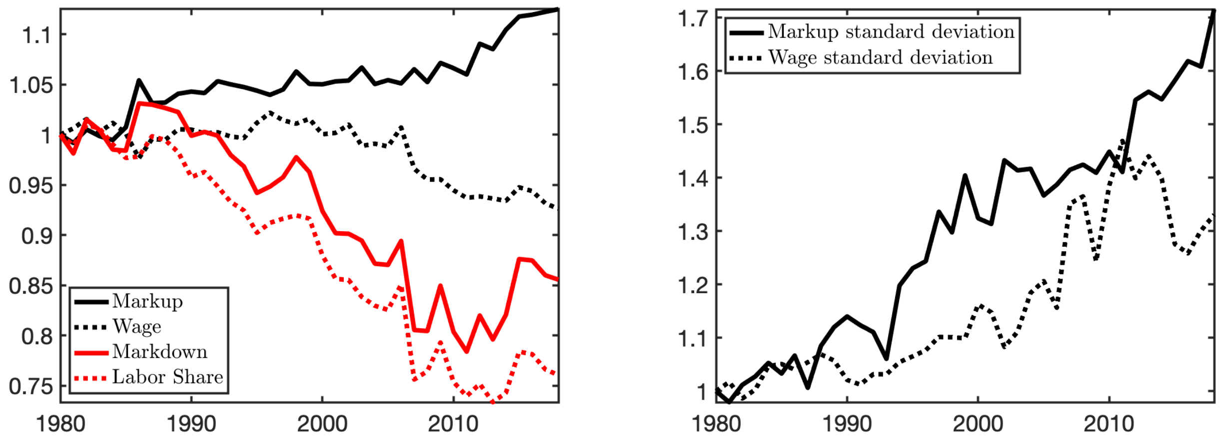
Notes: All parameters are time-variant.

The output elasticity of labor α is set at $2/3$ and the discount factor $\beta = 0.96$. I assume that the period is a year and treat every year as a steady-state to match the moments every given year from 1980 to 2018. I normalize the entry cost κ to achieve a steady-state mass of firms $N = 1$. Since the level of the wage is uniquely pinned down by the lowest productivity z_{min}, I then calibrate z_{min} to match the aggregate wage in 1980 and leave it unchanged throughout the period. The firm exit rate δ is taken from the Longitudinal Business Database.

The remaining key parameters $\Theta = \{\xi, \sigma, \varepsilon, \lambda, b\}$ are calibrated jointly by the model. The Pareto tail ξ pins down the distribution of relative sales of firms, and is chosen to match the standard deviation of relative sales in each 6-digit industry. In order to eliminate industry variation, I define relative sales as the sales of a firm in an industry relative to the average sales of all firms in that industry. The aggregate markup is pinned down by the average elasticity of demand σ, which I calibrate to match the Compustat cost-weighted average markup. The key parameter that affects the markup dispersion is ε/σ. I use the standard deviation of markup to determine this parameter. Two remaining labor market parameters λ and b pin down the sales-weighted markdown and replacement rate. The replacement rate is defined as average weekly benefits over average weekly wage, which is made available taken from the Department of Labor. I drop some firms with estimated markups below one, which the model cannot generate. Here, I only report parameter choices in Table 3 for 1980, 1990, 2000, 2010, and 2018.

Model Performance. Figure 6 reports the model results that correspond to the empirical results in Figure 1. All the values are shown relative to the 1980 values. The solid lines represent the

Figure 6 The Evolution of Markup, Wage, Labor Share, and Markdown (Model)



(a) Average

(b) Standard Deviation

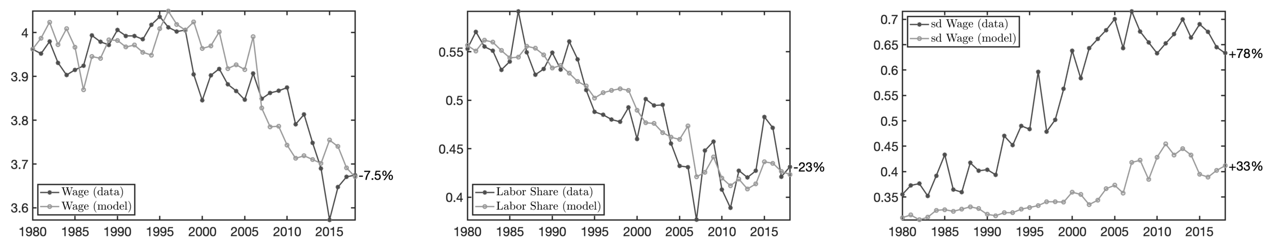
targeted moments that I match year by year, including the cost-weighted aggregate markup, markup standard deviation, and sales-weighted aggregate markdown. The dashed line represents the untargeted moments, including the employment-weighted aggregate wage, wage standard deviation, and the sales-weighted aggregate labor share.

The model perfectly matches the targeted moments.[22] The aggregate markup increased from 1.18 to 1.33; whereas, the aggregate markdown decreased from 0.99 to 0.85 between 1980 and 2018. As mentioned in Section 3, both increasing the aggregate markup and decreasing the aggregate markdown indicate that firms extract higher market power in both the product and labor markets. The dispersion in markup rose as in De Loecker *et al.* (2020) and a similar trend is observed for the standard deviation of relative sales. Hornstein *et al.* (2011) report the ratio of the mean to the minimum wage (Mm-ratio) as a summary statistic to compare wage dispersion across different classes of search models. This ratio is calibrated to match the replacement rate. I use the Department of Labor data to compute the ratio between average weekly benefit and average weekly wage and match the moment. Shimer (2005) and Tjaden & Wellschmied (2014) set the replacement rate at 40%; the replacement rate from the Department of Labor is 0.38 in 1980, and then it falls to 0.33 in 2018.

Figure 7 shows the comparison between model moments and empirical counterparts for the aggregate wage, labor share, and wage standard deviation. The model can produce the phenomenon that the aggregate wage increases slightly before 2000 and then decreases sharply after 2000. Similarly, the wage standard deviation increased throughout the time period, both in the model and the data, especially from 1990 to 2010. Finally, as the figures show, the de-

[22] I list 1980, 1990, 2000, 2010, and 2018 values in Table C.1

Figure 7 Untarget Moment Comparison



(a) Wage | (b) Labor Share | (c) Wage Standard Deviation

cline in the aggregate wage is associated with a decline in the labor share; thus, the model can generally reproduce the trends that show the decline in the aggregate wage, the decline in the labor share, and rising wage dispersion.

Labor Share. The aggregate labor share dropped roughly 23% in both the model and the data from 1980 to 2018. In the model, the decline in the aggregate labor share was mainly due to the increase in market powers in both the product and labor markets over time. There are other contributions to the decline in labor share[23]; however, the model is focused on the effects of the market powers on the labor shares only. Following Autor *et al.* (2017b), I can rewrite the aggregate labor share as

$$\underbrace{LS_t}_{\substack{\text{aggregate}\\ \text{labor share}}} = \sum_i \mathcal{SW}_{it} \times ls_{it} = \underbrace{\bar{LS}}_{\substack{\text{within}\\ \text{component}}} + \underbrace{cov\left(ls_{it}, \mathcal{SW}_{it}\right)}_{\substack{\text{between}\\ \text{component}}}$$

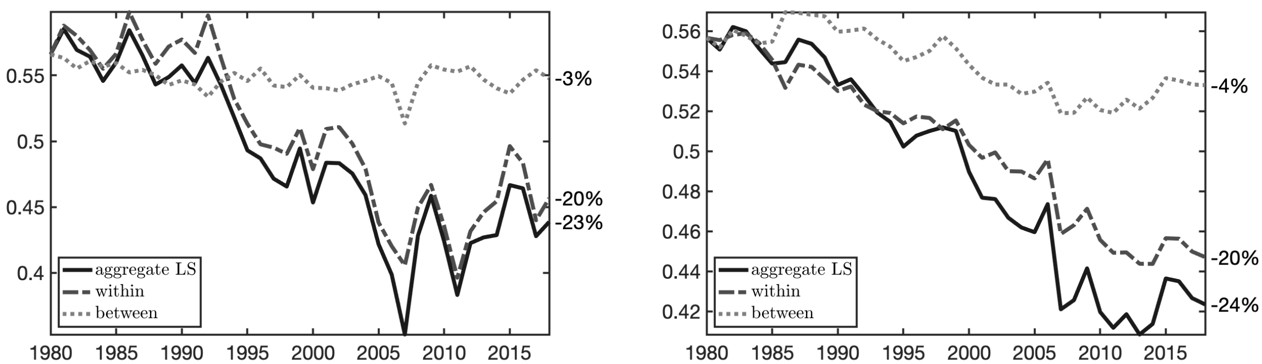
The aggregate labor share can be expressed as the sum of the average (unweighted) firm-level labor share (within-firm component) and the covariance between the labor share and the value-added share (between-firm component), which is the reallocation effect on the aggregate labor share. I compute both the within-firm component and between-firm component in the data and in the model. In the data, I compute each term separately in each 6-digit NAICS industry and then take a value-added weighted average.

Figure 8 shows the computed results both in the data and the model. Within 6-digit industries, the within-firm component contributes the most to the decline in labor share.[24] Table 4

[23]See Karabarbounis & Neiman (2013), Elsby *et al.* (2013), Autor *et al.* (2017a) and Kehrig & Vincent (2017)

[24]Although some literature finds that the between-firm component is the important source of the decline in the aggregate labor share (Autor *et al.*, 2017b; Kehrig & Vincent, 2017), the results here are not contradicting previous works. Most resource reallocation occurs at the sector level, e.g., 2-digit NAICS industry; however, I compute the components within the industry level. The purpose of this paper is to understand the effect of firm-level market power on a firm's behavior to adjust its labor share and wage. It provides another point of view to understand

Figure 8 Labor Share Decomposition



(a) Data

(b) Model

Table 4 Labor share decomposition of 10-year change

			80-90	90-00	00-10	10-18
LS	aggregate labor share	Data	0.56	0.51	0.44	0.43
		Model	0.55	0.51	0.46	0.42
			average cumulative change (%)			
ΔLS	change in labor share		-0.4	-6.3	-12.7	-13.4
			-0.6	-4.4	-10.4	-13.5
$\Delta \bar{LS}$	change in within		+0.8	-4.2	-10.5	-11.5
			-1.1	-3.9	-7.6	-10.7
$\Delta cov(ls, \mathcal{SW})$	change in between		-1.2	-2.1	-2.3	-1.8
			+0.5	-0.4	-2.8	-2.8

Notes: Row (1) and (2) show the average labor share from data and model in 10 years, e.g. 1980s, 1990s, 2000s, and 2010s. Row (3) to (8) show the average cumulative change for change in total, change in within, and change in between from 1980 to 1990, 2000, and 2018.

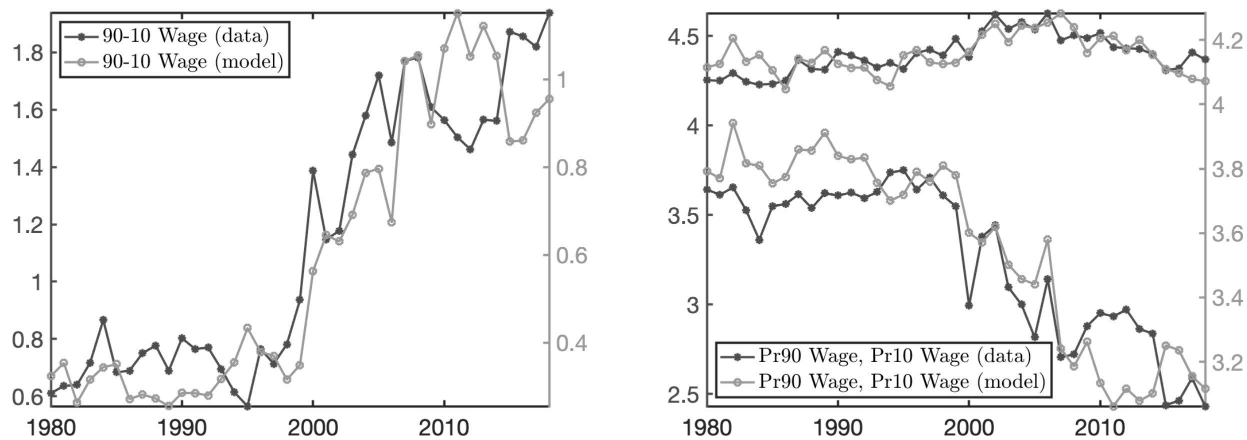
considers the cumulative change in the labor share over 10-year periods. The decomposition results show that 80% of the change in the labor share comes from the change in the within-firm component. The remaining 20% comes from the between-firm component, which is relatively small.

Wage and Dispersion. Although the model does not capture the exact point when the wage started to drop as shown in the data, the general trends of the aggregate wage are roughly the same. Both the data and model results show that the aggregate wage falls by 7.5% in 2018

the salient decline in the aggregate labor share.

relative to the 1980 level. On the other hand, the model fails to match the exact *level* of wage standard deviation and the *magnitude* of the increase in the wage standard deviation. The model shows that the wage standard deviation is 0.31 in 1980 and would increase by 33% to 0.41. However, the increase in the wage standard deviation is roughly 78% in data—from 0.35 to 0.67. The model can explain half of the wage dispersion. The contribution of increasing wage standard deviation in the model is mainly through the decline in the aggregate markdown and replacement rate. Various sources can contribute to wage dispersion. There is a difference between the model and the data because the model treats workers as homogeneous. An increasing number of studies have documented that the skilled-based technology is one of the main drivers of wage differential (Trottner, 2019). However, this paper is focused on the contributions of market power to wage dispersion and gives the different aspects of its dynamics.

Figure 9 Wage Dispersion



(a) Interdecile Range of Log Wage

(b) 90 and 10 Percentile Log Wage

Another way to examine wage dispersion is through the 90-to-10 ratio, that is, the interdecile range of wages. Moreover, comparing the trends in the 90 percentile of wages and the 10 percentile of wages allows me to investigate the source of the wage dispersion. Figure 9 plots the trends in interdecile range of employment-weighted wages, the 90 percentile of employment-weighted wages, and the 10 percentile of employment-weighted wages. The corresponding values are shown in Table 5. The interdecile range of wages has increased sharply, from 0.6 in 1980 to nearly triple that in 2018. The sharp increase mainly occurred between 1990 and 2010. Similar to the standard deviation of wage, the model underpredicts the interdecile range of wage; however, the trend matches. Similarly, the sharp rise in wage dispersion takes place between 1990 and 2010. Most importantly, the rising wage dispersion primarily

Table 5 The evolution of wage dispersion (90-to-10 ratio)

			1980	1990	2000	2010	2018
$W_{90} - W_{10}$	employment-weighted	Data	0.61	0.80	1.39	1.56	1.94
	interdecile range of log wage	Model	0.32	0.29	0.56	1.07	0.96
W_{90}	employment-weighted		4.25	4.41	4.38	4.52	4.37
	90 percentile log wage		4.12	4.13	4.16	4.21	4.07
W_{10}	employment-weighted		3.64	3.61	2.99	2.95	2.43
	10 percentile log wage		3.79	3.84	3.60	3.13	3.12

Notes: All values list in corresponding years.

stems from the decline in low-percentile wage. Throughout the time period, the 10 percentile of wages has dropped 33% in the data and 18% in the model, whereas the change in the 90 percentile of wages was negligible. Thus, the model can capture the fact that it is the low percentile of the wage distribution that pulls down the aggregate wage and it occurs around 2000.

5 The Consequences of Market Power

In this section, I start with my primary analysis: How does market power affect the labor share and wage dispersion? I will discuss the three main group of parameters, that is, demand elasticity, labor market competitiveness, and productivity dispersion, all of which affect firm-level markup and markdown and their contributions to the labor share and wage dispersion. Next, I use the calibrated model to simulate the effects of each group of parameters. Finally, I use the model simulation to run the regression analysis to test the effects of firm-level markup and markdown on the labor share and wage dispersion.

Counterfactual. In the model, three main sources affect the labor share, aggregate wage, and wage dispersion. Sizable prior literature has already shown the implication of the rising markup on the decline in the labor share.[25] Nevertheless, few studies directly explore the link between product market power and wage.[26] The parameters that govern the aggregate markup and markup dispersion are the demand elasticity, σ, and superelasticity, ϵ/σ. Therefore, I first consider only changing these parameters to analyze the effect of markups on the labor share and wages. Berger *et al.* (2019) show that the labor market concentration makes an important contribution to the decline in the labor share. Several papers have documented the

[25] See De Loecker *et al.* (2020), Baqaee & Farhi (2017), and Edmond *et al.* (2018)

[26] Coşar *et al.* (2016) inspects the effect of trade integration with global product markets on wage inequality

impact of labor market power on wages and its dispersion.[27] The parameters that govern the aggregate markdown are labor market competitiveness, $k = \lambda/\delta$, and unemployment benefit, b. I change these parameters to understand how markdowns change the aggregate labor share and wages. Lastly, Gouin-Bonenfant *et al.* (2018) find that a rise in the dispersion of firm productivity causes the aggregate labor share to decline as productivity dispersion effectively shields high-productivity firms from wage competition.[28] Therefore, I change the productivity dispersion parameter, ξ, to observe its contribution to the labor share and wages.

Role of Demand Elasticity. Estimating demand elasticity requires detailed price and physical quantity data as in Foster *et al.* (2008). Although Compustat does not contain such detailed information, I can still compute the aggregate demand elasticity by using the implications of profit maximization. The optimal price-setting behavior implies that the markup over marginal cost is equal to $\frac{\sigma}{\sigma-1}$. Therefore, I can back up the aggregate demand elasticity by the inverse of the aggregate markup, which decreased from 5.56 in 1980 to 3.03 in 2018. The calibrated demand elasticity from the model (shown in Figure 10 (a)) decreased from 8.98 to 4.82, which is above the simple inverse markup approach.[29] In fact, the model calibrates higher demand elasticity because the model features variable markup. Traditional macro literature assumes the demand elasticity is 10, but the trade literature assumes it to be around 4. There is still no clear way to estimate the aggregate demand elasticity, but I view the calibration result as close to that found in the literature. Most of the increasing markup is a result of the decline in the aggregate demand elasticity.[30] Figure 10 (b) shown the markup trend when I allow only the demand elasticity parameters to change over time. If only the demand elasticity parameters decline, the aggregate markup rises from 1.18 to 1.30, which is not much different from the calibrated moment that rose from 1.18 to 1.33. Therefore, the demand elasticity parameters are the main channel that leads to an increase in the aggregate markup; it allows me to undertake a counterfactual analysis to understand the consequences of product market power on the labor share and wages.

Role of Labor Market Competitiveness. The calibrated job arrival rates (job-finding rates) are lower than 35-40% of that found in the literature. The model calibrates low job arrival rates

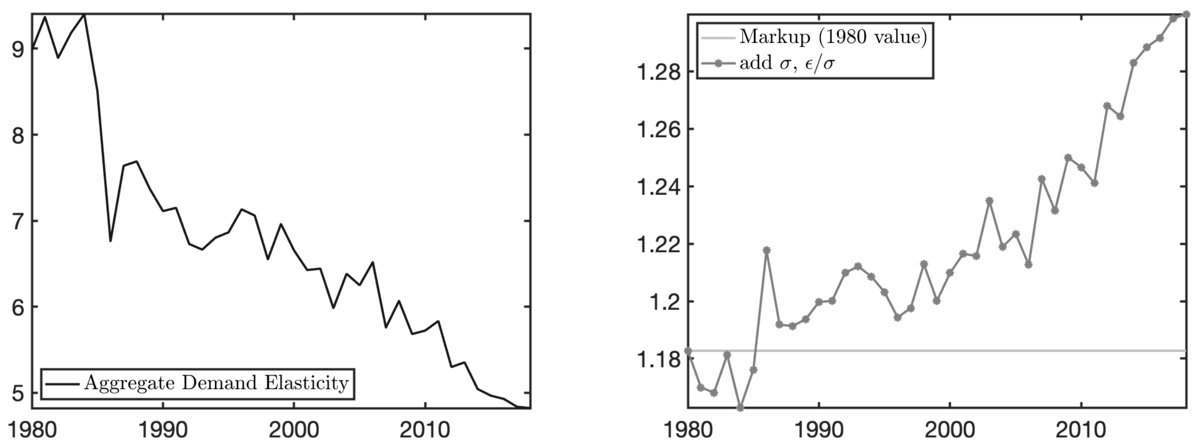
[27]See Lentz & Mortensen (2010), Alan (2011), and Song *et al.* (2019).

[28]The link between productivity and wage dispersion is also documented in Dunne *et al.* (2004), Faggio *et al.* (2010), and Barth *et al.* (2016). On the other hand, Kehrig & Vincent (2021) documents that the reallocation of value-added to the lower labor share distribution is the cause of the decline in the aggregate labor share.

[29]Oberfield & Raval (2014) use detail plant-level data, the U.S. Census of Manufactures and Annual Survey of Manufactures (ASM), to estimate industry-wide demand elasticities using a similar inverse markup approach and obtain the range 3 to 8.

[30]One of the implications of the decline in aggregate demand elasticity is household spending concentration or consumer inertia, as in Gourio & Rudanko (2014), Neiman *et al.* (2018), and Bornstein *et al.* (2018)

Figure 10 Aggregate Demand Elasticity and Aggregate Markup

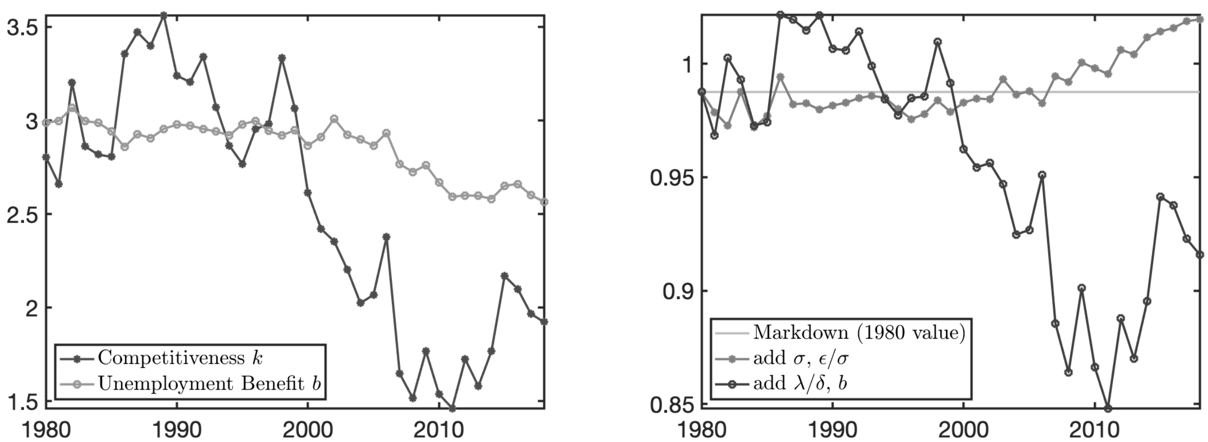


(a) Aggregate Demand Elasticity

(b) Markup

simply because the job destruction rate I obtain from the Longitudinal Business Database (LBD) is lower than the traditional value of 15%. There are two measurements for job destruction rate in the LBD. One only considers that jobs are destroyed by the death of the establishment, which the exiting firm creates. The other one includes jobs destruction by existing firms. The average job destruction rate from the exiting firms is 5%, whereas from both the exiting and the existing firms is 15%. When I use the average job destruction rate from both the exiting and the existing firms to calibrate the model, the job arrival rates would be three times the values shown in Table 1. Thus, the relevant parameter governing labor market competitiveness is the ratio of the job arrival rate and job destruction rate, $k = \lambda/\delta$. A higher value of k indicates more competition among firms for workers, which is expected to raise wages. Therefore, a higher value of k represents the labor market being closer to a competitive market; whereas, a lower value of k indicates a noncompetitive labor market. Figure 11 (a) plots model calibrated value k and the reservation wage (unemployment benefit) b. Labor market competitiveness has dropped dramatically between 1990 and 2010—the same period when the aggregate markdown also declined. The unemployment benefit gradually declines over time since the replacement rate dropped from 0.38 to 0.33. The effect of imperfect competition in the labor market is the main driver of decreasing the aggregate markdown. As shown in Figure 11 (b), there is a clear downward trend in the aggregate markdown between 1990 and 2010 once I further allow the labor market parameters to change over time. The graph also shows the markdown trend if I only change the demand elasticity parameters. Apparently, most of the decline in the aggregate markdown results from the decline in labor market competitiveness, specifically as the labor market became less competitive. Therefore, I analyze the effect of markdowns on the labor share and wages by changing the labor market competitiveness parameters.

Figure 11 Labor Market Competitiveness and Markdown



(a) Competitiveness & Unemployment Benefit

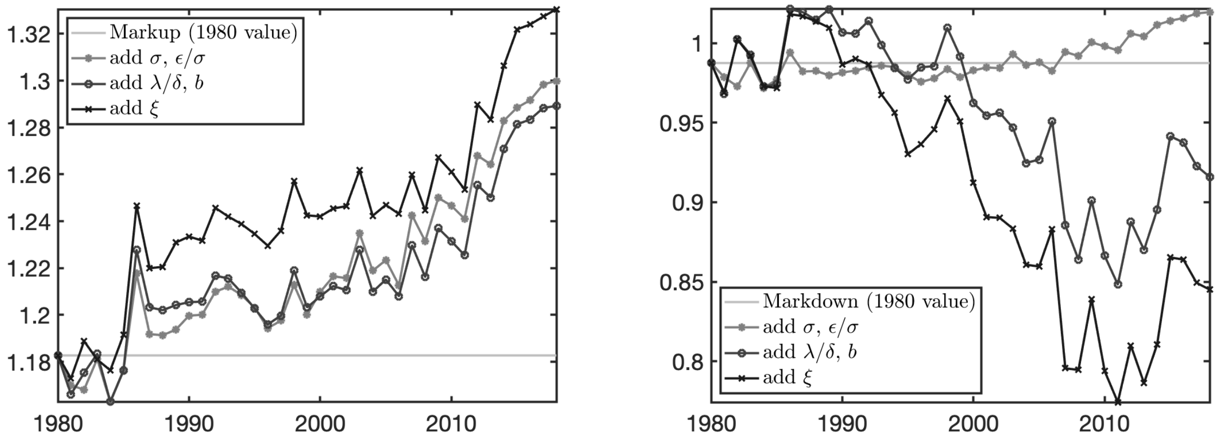
(b) Markdown

Role of Productivity Dispersion. The calibration shown in Table 1 indicates that the Pareto tail decreased from 6.12 in 1980 to 4.04 in 2018, which is close to Arkolakis *et al.* (2019), while the calibrated Pareto tail is 5 using French firm-level data. The decline in the Pareto tail means the productivity dispersion has increased throughout time. The dispersion in productivity plays an important role in increasing markup and decreasing markdown in the model as well. There are several empirical works that explore changes in productivity dispersion and its implications, such as Syverson (2004) and Aghion *et al.* (2009). The calibration results in this paper are also in line with the finding that the productivity dispersion has increased over time. Figure 12 show both markup and markdown trends after I allow productivity dispersion to increase over time. Productivity dispersion can also explain part of the rise in the aggregate markup and the decline in the aggregate markdown. Increasing productivity dispersion has reallocation effects on firm-level sales; therefore, resource are reallocated to high-productivity firms, which feature a higher markup and a lower markdown. Hence, the rise in markup and the fall in markdown are affected by productivity dispersion through the channel of resource reallocation.[31] Therefore, I can also understand the impact of productivity dispersion on the labor share and wages.

Counterfactual Results. Figure 13 plots the counterfactual cases for the moments that I am interested in: the aggregate wage, aggregate labor share, and wage standard deviation. Each

[31] De Loecker *et al.* (2020) documents the same reallocation effect.

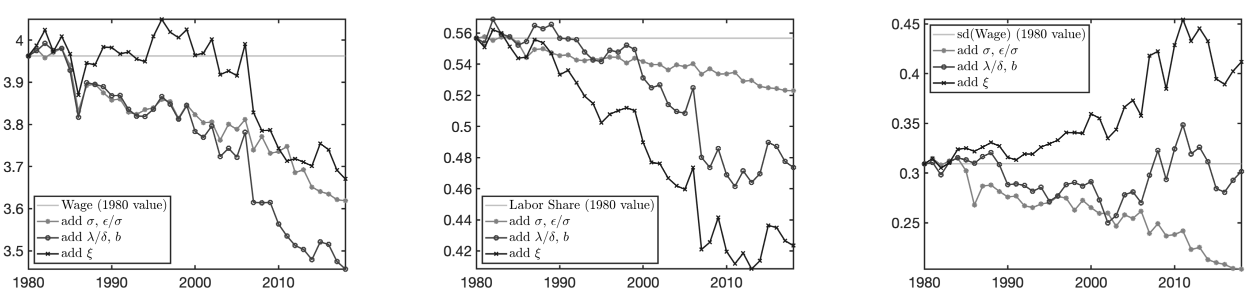
Figure 12 Product Market Power and Labor Market Power



(a) Markup (b) Markdown

group of parameters play a different role in affecting each outcome variable. The purpose of the counterfactual analysis is to isolate the contributions of each different factor to the aggregate labor share, aggregate wage, and wage dispersion. Table 6 lists the percentage change from 1980 to 1990, 2000, 2010, and 2018 in each different counterfactual case for five variables of interest, respectively. The values in each line represent the percentage change from 1980 to the specific year. Case in point, the aggregate wage rose 0.5% from 1980 to 1990, 0.04% to 2000, and fell 5.53% to 2010 and 7.36% to 2018. Therefore, I can see the aggregate wage fell the most after 2000.

Figure 13 Counterfactual Labor Share, Wage, and Wage Standard Deviation



(a) Wage (b) Labor Share (c) Wage Standard Deviation

The lines for *markup*, *markdown*, and *productivity* indicate the effect of that particular channel on the variables. The *markup* lines show the counterfactual case when I allow only the demand elasticity parameters to change over time. I then add the change in labor market parameters in the *markdown* lines. Lastly, the effects of change in productivity dispersion are shown in the *productivity* lines. The values of *markup*, *markdown*, and *productivity* shown below each specific variable in a given year represent the effect of the specific channel on the variable of interest. For instance, in 2018, the *markup* channel caused the aggregate wage to fall by 8.67%, the *markdown* channel added a further decline of 4.09%, and the *productivity* channel increased aggregate wages by around 5.39%. Therefore, aggregating all the channels, I obtain the aggregate wage decline of 7.36% from 1980 to 2018.

Counterfactual Wage. The aggregate wage started to increase from 1980 to 1990 then gradually declined. Until 2018, the aggregate wage dropped 7.36%. In Figure 13 (a), it can be seen that, had only demand elasticity parameters declined, the aggregate wage would have dropped to around 3.6 in 2018, which is more than the actual aggregate wage drop. Furthermore, the aggregate wage would have declined to less than 3.5 in 2018 if I further allowed the labor market parameters to change. Most importantly, the further decline in wages started in 2000. These two counterfactual analyses indicate that product market power causes the aggregate wage to decline over time and labor market power exacerbates it after 2000. Increasing productivity dispersion brings the aggregate wage to a higher level since high-productivity firms became more productive and have a higher marginal productivity of labor, and a higher marginal product of labor would raise the wages they pay. As a result, the aggregate wage was stable before 2000 due to the effect of productivity dispersion offsetting the markup effect; whereas, the aggregate wage started to fall after 2000 since the markdown effect became another source of decline in the aggregate wage.

Table C.2 also shows the other aggregate variables that might help explain the effects from each channel. Both the *markup* channel and the *markdown* channel cause the aggregate output to decline and the *markdown* channel becomes more prominent after 2000. Similar trends can be found in aggregate labor income. The large drop in the *markdown* effect mostly came from the decline in production labor. As shown in equation (25), aggregate output is associated with aggregate productivity and production labor. The effect from of the *markup* channel on wages works through mainly by reducing the aggregate productivity; however, the *markdown* channel slightly increased the aggregate productivity but significantly reduced production labors. Therefore, the effect from the *markdown* channel not only directly reduced the aggregate wage but also indirectly brought down the number of workers and decreased the aggregate output. This labor resource reallocation is mainly the result of the drop in labor market competitiveness. To be more specific, the effect of the *markup* on wages comes from firms reducing the wage they pay, but the effect from the *markdown* on wages results from reducing the size of employment and shielding away from wage competition. In sum, the markup has an impact on the decline in the aggregate wage, but the markdown contributes to it after 2000.

Table 6 Counterfactual

		1990	2000	2010	2018
		percentage change (%)			
W	employment-weighted aggregate wage	+0.51	+0.04	-5.53	-7.36
	markup	-2.64	-3.52	-5.72	-8.67
	markdown	+0.27	-1.01	-4.34	-4.09
	productivity	+2.88	+4.57	+4.53	+5.39
$sd(W)$	wage standard deviation	+2.09	+16.23	+38.52	+33.06
	markup	-10.77	-14.28	-22.99	-34.19
	markdown	+3.95	+8.50	+27.78	+31.69
	productivity	+8.92	+22.01	+33.74	+35.56
LS	sales-weighted aggregate labor share	-4.22	-12.02	-24.62	-23.93
	markup	-2.01	-2.70	-4.11	-6.06
	markdown	+2.03	-1.87	-11.64	-8.85
	productivity	-4.24	-7.44	-8.86	-9.02
W_{90}	employment-weighted 90 percentile wage	+0.25	+1.17	+2.19	-1.07
	markup	-2.63	-3.50	-5.66	-8.51
	markdown	+1.36	-1.41	-1.85	-3.72
	productivity	+1.53	+6.07	+9.70	+11.16
W_{10}	employment-weighted 10 percentile wage	+1.27	-5.05	-17.33	-17.81
	markup	-1.71	-2.28	-3.71	-6.64
	markdown	+1.04	-3.24	-13.92	-11.59
	productivity	+1.95	+0.47	+0.29	+0.42

Notes: All values shows the percentage change from 1980 to 1990, 2000, and 2018.

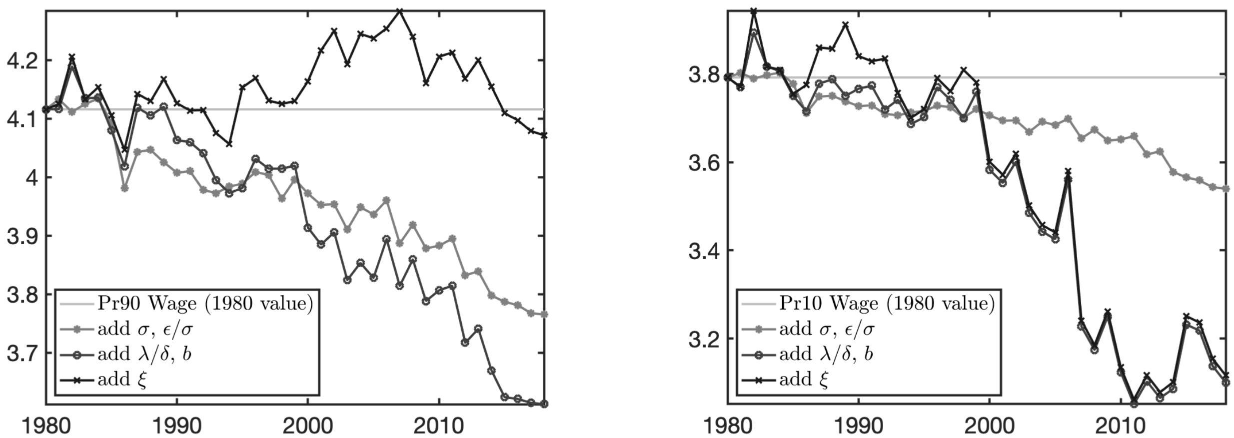
Counterfactual Labor Share. Before 1990, there was a slight drop in the labor share—around 4%. The rapid decline in the labor share took place between 1990 and 2010, then leveled off afterward. In Figure 13 (b), both demand elasticity parameters and productivity dispersion explain the decline in the labor share throughout the time period. The *markup* channel contributes roughly to one-fourth of the drop and *productivity* explains around 40% of the decline in the labor share. From equation (27), the aggregate labor share is associated with labor income (WL) and aggregate output (Y). As mentioned earlier, the *markup* channel affects the labor share mostly via a fall in wages, with a slight reduction in aggregate output. However, productivity dispersion increases the aggregate output the most; hence, it reduces the aggregate labor share. Although productivity dispersion also increases labor income, the effect on labor income is disproportional to the effect on aggregate output, the increase in labor income is only about one-third of the aggregate output increase. Therefore, a rise in productivity dispersion would result in a decline in the aggregate labor share.

Furthermore, the *markdown* channel reduces the aggregate labor share the most after 2000. Again, most of the effect came from a sharp decline in labor income. Therefore, the *markdown* effect can explain 35% of the drop in the aggregate labor share between 1980 and 2018. Table C.3 also shows the cumulative change in the labor share decomposition over 10-year periods. Each of the three channels can explain the within-firm component of the decline in the labor share and the *markdown* and *productivity* channels can explain most of the between-firm component since both channels relate to the resource reallocation in the product market and labor market. This section has provided a brief summary of the literature relating to the decline in the aggregate labor share. The evidence reviewed here suggests that each channel plays an important role in explaining the decline in the aggregate labor share.

Counterfactual Wage Dispersion. The standard deviation of wages rapidly increased between 1990 and 2010. Based on Figure 13 (c), almost half of the increase in wage dispersion can be explained by the decline in the aggregate markdown. Both labor market competitiveness and unemployment benefits parameters help explain both the decline in the aggregate markdown and the rise in wage dispersion,[32] especially after 2000. On the other hand, allowing firms to exploit markup would decrease the wage dispersion because every firm pays a lower wage. Similar to the literature, increasing productivity dispersion by itself has a significant impact on wage dispersion. A large dispersion in productivity means a large dispersion in the marginal product of labor; therefore, the wage differential becomes more dispersed. The *productivity* channel helps improve market competition by increasing the number of firms and new entrants; in turn, it increases the aggregate output and aggregate productivity. Higher productivity dispersion allows larger firms to reach higher productivity levels and pay higher wages. Therefore, the *productivity* channel explains the other half of wage dispersions.

[32] Hornstein *et al.* (2011) discussed, for a large class of search models, that wage dispersion can be explained by the mean-min wage ratio, which is the concept of the replacement rate in the model, and I calibrated the unemployment benefits to match the replacement rate.

Figure 14 Counterfactual Wage Percentile



(a) 90 Percentile Wage

(b) 10 Percentile Wage

Another way to examine wage dispersion is through the employment-weighted interdecile range of wages. Figure 14 (a) and (b) show the 90 and 10 percentile of wages with three counterfactual analyses over the sample period. The employment-weighted 90 percentile wage has increased around 2% until 2010, and then declined to -1%; whereas the employment-weighted 10 percentile wage has dropped dramatically between 2000 and 2010 (around 17%). Both the 90 and 10 percentile wages would have dropped significantly if only the demand elasticity parameters changed over time. This shows that the *markup* channel causes both the high and low percentile of wages to decline, but the decline is larger in high percentile of wages than in low percentile of wages. As a result, it slightly closed the gap between the 90 and 10 percentile wages, and, in turn, the wage dispersion decreased. Hence, it could be hypothesized that the markup, or product market power, caused the decline in both wages and wage dispersion.

Both the effects of the *productivity* and the *markdown* channels on wages contribute to the rise in wage dispersion, although they work through different mechanisms. As mentioned earlier, the low-productivity firm finds it more profitable to extract monopsony power and lower the wage it pays. Consequently, around two-third of the decline in 10 percentile wages come from the *markdown* effect. On the other hand, large productivity dispersion allows large firms to increase their productivity the most and pay a large wage premium. Because of this, the *productivity* channel improves the high-percentile wage, which almost offsets the effects of *markups* and *markdowns*; hence, the 90 percentile wage slightly increased in the past four decades.

In sum, the *markup* channel has significant impact on the decline in wage for both high and low percentiles and in wage dispersion. Both *markdown* and *productivity* channels explain the rise in wage dispersion caused by a reduction in the low percentile wage from *markdown* effect and an increase in the high percentile wage from *productivity* dispersion effect.

Regression Analysis. To test the model validity to the effect of market power on labor share and wage dispersion, I run the same regression analysis as shown in section 2. In Table 7, I report the regression coefficients of the log of the labor share on the log of the firm's markup and markdown; all columns are clustered at the firm level. The first three specifications are reported in Table 1 (columns four to six). The last two specifications show the regression results from the model simulation. Although the coefficients of markdown are larger in the model than in the data, the effect of the markup is quite similar—around -0.87 to -0.91. The difference in the markdown coefficient between the model and the data mostly results from the mismatch in the wage distribution. The model does not consider any other source that can affect the wage firms pay, and firms pay different wages because they face different labor supply elasticities and charge different markdowns. If the skill-premium is also an important determinant in wage payment decisions, the effect from the markdown would be smaller, and this is the exact mechanism that Trottner (2019) proposed. This paper plays a complementary role to the vast literature on the labor share and wage dispersion, and it provides one reason for the salient fact of the decline in the labor share. In summary, the model can accurately predict the casual effect of the markup and markdown to the labor share.

Table 7 Data v.s. Model: The effects of markups and markdowns on labor shares

	(1) Data	(2) Data	(3) Data	(4) Model	(5) Model
Markup (log)	-0.877***	-0.847***	-0.906***	-0.915***	-0.874***
	(0.0602)	(0.0609)	(0.0692)	(0.0622)	(0.0506)
Markdown (log)	0.399***	0.387***	0.375***	1.073***	1.065***
	(0.0247)	(0.0246)	(0.0207)	(0.0096)	(0.0100)
Year F.E.		X	X		X
Industry F.E.			X		
Observations	23282	23282	23282		
R^2	0.148	0.155	0.570	0.431	0.432

Lastly, I test the prediction that the heterogeneous effects of markup and markdown lead to a decline of the aggregate wage in the model. In Table 8, the first-three columns report the results in Table 2. Similar heterogeneous effects can be found in the model simulation (columns four to six in Table 8); the effects of markups or markdowns from upper-wage percentile firms

Table 8 Data v.s. Model: The effects of markups and markdowns on wages

	(1) Data W	(2) Data $< W_{25}$	(3) Data $> W_{75}$	(4) Model W	(5) Model $< W_{25}$	(6) Model $> W_{75}$
Markup (log)	-0.855***	-0.885***	-0.456***	-0.990***	-0.448***	-0.288***
	(0.0402)	(0.0665)	(0.0517)	(0.0124)	(0.0316)	(0.0151)
Markdown (log)	0.875***	0.828***	0.505***	1.005***	0.479***	0.227***
	(0.0122)	(0.0191)	(0.0304)	(0.0095)	(0.0159)	(0.0135)
Employment (log)	-0.871***	-0.786***	-0.551***	-0.993***	-0.481***	-0.275***
	(0.0116)	(0.0184)	(0.0293)	(0.0116)	(0.0280)	(0.0140)
Sales (log)	0.854***	0.772***	0.525***	0.990***	0.464***	0.290***
	(0.0112)	(0.0176)	(0.0287)	(0.0119)	(0.0325)	(0.0143)
Year F.E.	X	X	X	X	X	X
Industry F.E.	X	X	X			
Observations	23368	5858	5772			
R^2	0.871	0.867	0.825	0.581	0.427	0.267

are smaller than the effects from lower-wage percentile firms. Although the magnitude of the effect is smaller than in the data, the model can produce quite well the heterogeneous effects on the upper and lower-wage percentiles.

6 Conclusion

In this paper, I study the role of firms' market power in both the product market and labor market as drivers of the decline in the aggregate labor share and the aggregate wage after 1980. I argue that the prevalence of the rising aggregate markup and the falling aggregate markdown play important roles in these two salient facts. Specifically, the decline in the aggregate wage can be attributed mostly to the lower percentile of the wage distribution. As a result, wages in the left tail of the distribution pull down aggregate wages and drive up wage dispersion.

Firms' market power rose not just in the product market but also in the labor market over the past four decades. I document that the aggregate markup rose around 12.5% and the aggregate markdown fell around 15% between 1980 and 2018. I then build a model that features heterogeneous firms engaged in monopolistic competition with non-CES demand with the Kimball (1995) aggregator and on-the-job search in the labor market to calibrate the rising market power in both markets. The model can reproduce both the decline in the aggregate labor share and the aggregate wage. Moreover, the model can predict the rise in wage dispersion that stems from a sharp decline in the lower wage percentiles after 2000. The increasing aggregate markup has a direct impact on the decline in the aggregate wage. In addition, there is a heterogeneous effect of markups on firms' monopsony power and wages. Although low-productivity firms claim less markup than high-productivity firms, low-productivity firms lower the wage more than high-productivity firms. I find that the sharp decline in the lower wage percentiles is caused by both the larger decline in the labor markdown from low-productivity firms and the decreasing replacement rate in the labor market. These two facts indicate that labor market competitiveness is an important channel for both the decline in the aggregate wage and the rise in wage dispersion.

In the counterfactual analysis, I consider three important channels for both increasing markup and decreasing markdown. The productivity dispersion channel can explain a rise in wage dispersion, which is caused by a rise in the higher wage percentiles; whereas, the markdown channel can rationalize the sharp decline in the lower wage percentiles. The markup channel has a negative impact on all wage levels. Therefore, the effects from markups and productivity dispersion offset each other at the higher wage percentiles but both markup and markdown contribute to the decline in the lower wage percentiles. As a consequence, the sharp decline in the aggregate wage after 2000 is the result of increasing firm monopoly power and monopsony power. Finally, the productivity dispersion explains the most of the decline in the labor share—up to 40%. The markup and markdown channels explain 25% and 35%, respectively. Increasing productivity dispersion would reallocate the resources to large firms. This resource allocation also contributes to the rise in markup and the fall in markdown. In sum, both product market power and labor market power are equally important in explaining the decline in the aggregate labor share.

The future research agenda could include, firstly, computing misallocation caused by variable markups and wage dispersion. Secondly, most of the increasing markup is a result of the decline in the aggregate demand elasticity in the model, future work could factor in the the demand-side driven factor. Detailed modeling of consumer behavior can rationalize the evolution of markups and its consequences for the labor share and wages. Lastly, adding heterogeneous firms' international trade specification ((Melitz, 2003; Melitz & Ottaviano, 2008)) to gauge how trade affects both variable markups and wages is also a promising research avenue because of globalization especially after 2000.

References

Ackerberg, D.A., Caves, K. & Frazer, G. (2015) Identification properties of recent production function estimators. *Econometrica* **83**, 2411–2451.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. & Prantl, S. (2009) The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics* **91**, 20–32.

Akcigit, U. & Ates, S.T. (2019) Ten facts on declining business dynamism and lessons from endogenous growth theory. Tech. rep., National Bureau of Economic Research.

Alan, M. (2011) Imperfect competition in the labor market. *Handbook of labor economics*, vol. 4, pp. 973–1041, Elsevier.

Arkolakis, C., Costinot, A., Donaldson, D. & Rodríguez-Clare, A. (2019) The elusive pro-competitive effects of trade. *The Review of Economic Studies* **86**, 46–80.

Autor, D., Dorn, D., Katz, L., Patterson, C. & Van Reenen, J. (2017a) Dp11810 concentrating on the fall of the labor share .

Autor, D., Dorn, D., Katz, L.F., Patterson, C., Van Reenen, J. *et al.* (2017b) *The fall of the labor share and the rise of superstar firms*. National Bureau of Economic Research.

Axtell, R.L. (2001) Zipf distribution of us firm sizes. *science* **293**, 1818–1820.

Baqaee, D.R. & Farhi, E. (2017) Productivity and misallocation in general equilibrium. Tech. rep., National Bureau of Economic Research.

Barth, E., Bryson, A., Davis, J.C. & Freeman, R. (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**, S67–S97.

Berger, D.W., Herkenhoff, K.F. & Mongey, S. (2019) Labor market power. Working Paper 25719, National Bureau of Economic Research.

Bond, S., Hashemi, A., Kaplan, G. & Zoch, P. (2020) Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. Tech. rep., National Bureau of Economic Research.

Bornstein, G. *et al.* (2018) Entry and profits in an aging economy: The role of consumer inertia. Tech. rep., mimeo.

Brown, C. & Medoff, J. (1989) The employer size-wage effect. *Journal of political Economy* **97**, 1027–1059.

Burdett, K. & Mortensen, D.T. (1998) Wage differentials, employer size, and unemployment. *International Economic Review* pp. 257–273.

Coles, M.G. & Mortensen, D.T. (2016) Equilibrium labor turnover, firm growth, and unemployment. *Econometrica* **84**, 347–363.

Coşar, A.K., Guner, N. & Tybout, J. (2016) Firm dynamics, job turnover, and wage distributions in an open economy. *American Economic Review* **106**, 625–63.

Davis, S.J., Haltiwanger, J., Jarmin, R., Miranda, J., Foote, C. & Nagypal, E. (2006) Volatility and dispersion in business growth rates: Publicly traded versus privately held firms [with comments and discussion]. *NBER macroeconomics annual* **21**, 107–179.

De Loecker, J. & Eeckhout, J. (2018) Global market power. Tech. rep., National Bureau of Economic Research.

De Loecker, J., Eeckhout, J. & Unger, G. (2020) The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics* .

De Loecker, J. & Warzynski, F. (2012) Markups and firm-level export status. *American economic*

review **102**, 2437–71.
Decker, R.A., Haltiwanger, J., Jarmin, R.S. & Miranda, J. (2016) Where has all the skewness gone? the decline in high-growth (young) firms in the us. *European Economic Review* **86**, 4–23.
Dunne, T., Foster, L., Haltiwanger, J. & Troske, K.R. (2004) Wage and productivity dispersion in united states manufacturing: The role of computer investment. *Journal of Labor Economics* **22**, 397–429.
Edmond, C., Midrigan, V. & Xu, D.Y. (2015) Competition, markups, and the gains from international trade. *American Economic Review* **105**, 3183–3221.
Edmond, C., Midrigan, V. & Xu, D.Y. (2018) How costly are markups? Tech. rep., National Bureau of Economic Research.
Elsby, M.W., Hobijn, B. & Şahin, A. (2013) The decline of the us labor share. *Brookings Papers on Economic Activity* **2013**, 1–63.
Faggio, G., Salvanes, K.G. & Van Reenen, J. (2010) The evolution of inequality in productivity and wages: panel data evidence. *Industrial and Corporate Change* **19**, 1919–1951.
Falch, T. (2010) The elasticity of labor supply at the establishment level. *Journal of Labor Economics* **28**, 237–266.
Flynn, Z., Gandhi, A. & Traina, J. (2019) Measuring markups with production data. Tech. rep., Working paper.
Foster, L., Haltiwanger, J. & Syverson, C. (2008) Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* **98**, 394–425.
Gabaix, X. (2011) The granular origins of aggregate fluctuations. *Econometrica* **79**, 733–772.
Gouin-Bonenfant, E. *et al.* (2018) Productivity dispersion, between-firm competition and the labor share. *2018 Meeting Papers*, 1171, Society for Economic Dynamics.
Gourio, F. & Rudanko, L. (2014) Customer capital. *Review of Economic Studies* **81**, 1102–1136.
Gutiérrez, G. & Philippon, T. (2017) Declining competition and investment in the us. Tech. rep., National Bureau of Economic Research.
Hall, R.E. (1988) The relation between price and marginal cost in us industry. *Journal of political Economy* **96**, 921–947.
Helpman, E., Itskhoki, O., Muendler, M.A. & Redding, S.J. (2017) Trade and inequality: From theory to estimation. *The Review of Economic Studies* **84**, 357–405.
Hornstein, A., Krusell, P. & Violante, G.L. (2011) Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review* **101**, 2873–98.
Karabarbounis, L. & Neiman, B. (2013) The global decline of the labor share. *The Quarterly journal of economics* **129**, 61–103.
Kehrig, M. & Vincent, N. (2017) Growing productivity without growing wages: The micro-level anatomy of the aggregate labor share decline .
Kehrig, M. & Vincent, N. (2021) The micro-level anatomy of the labor share decline. *The Quarterly Journal of Economics* **136**, 1031–1087.
Keller, W. & Yeaple, S.R. (2009) Multinational enterprises, international trade, and productivity growth: firm-level evidence from the united states. *The review of economics and statistics* **91**, 821–831.
Kimball, M.S. (1995) The quantitative analytics of the basic neomonetarist model. Tech. rep., National Bureau of Economic Research.
Klenow, P.J. & Willis, J.L. (2016) Real rigidities and nominal price changes. *Economica* **83**, 443–472.

Lentz, R. & Mortensen, D.T. (2010) Labor market models of worker and firm heterogeneity. *Annu. Rev. Econ.* **2**, 577–602.

Levinsohn, J. & Petrin, A. (2003) Estimating production functions using inputs to control for unobservables. *The review of economic studies* **70**, 317–341.

Melitz, M.J. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* **71**, 1695–1725.

Melitz, M.J. & Ottaviano, G.I. (2008) Market size, trade, and productivity. *The review of economic studies* **75**, 295–316.

Neiman, B., Vavra, J. *et al.* (2018) The rise in household spending concentration. *Meeting Papers Society for Economic Dynamics*.

Oberfield, E. & Raval, D. (2014) Micro data and macro technology. Tech. rep., National Bureau of Economic Research.

Ransom, M.R. & Sims, D.P. (2010) Estimating the firm's labor supply curve in a "new monopsony" framework: Schoolteachers in missouri. *Journal of Labor Economics* **28**, 331–355.

Shimer, R. (2005) The cyclical behavior of equilibrium unemployment and vacancies. *American economic review* **95**, 25–49.

Song, J., Price, D.J., Guvenen, F., Bloom, N. & Von Wachter, T. (2019) Firming up inequality. *The Quarterly journal of economics* **134**, 1–50.

Staiger, D.O., Spetz, J. & Phibbs, C.S. (2010) Is there monopsony in the labor market? evidence from a natural experiment. *Journal of Labor Economics* **28**, 211–236.

Syverson, C. (2004) Product substitutability and productivity dispersion. *Review of Economics and Statistics* **86**, 534–550.

Tjaden, V. & Wellschmied, F. (2014) Quantifying the contribution of search to wage inequality. *American Economic Journal: Macroeconomics* **6**, 134–61.

Trottner, F. (2019) Who gains from scale? trade and wage inequality within and between firms. Tech. rep., Working paper.

Wooldridge, J.M. (2009) On estimating firm-level production functions using proxy variables to control for unobservables. *Economics letters* **104**, 112–114.

Appendices

A Estimation

Data. Following the definition in De Loecker *et al.* (2020) closely, I can observe firm sales, and I use "Cost of Goods Sold" (COGS) as the variable inputs costs, both variables are deflated by GDP deflator. I use "Property, Plant and Equipment" (PPEGT) as Gross Capital, which is adjusted for the industry-level input price deflator. I use variable XLR divided by "Employees" (emp) to compute firm-level wages. I further remove year and (6-digit NAICS) industry fixed effects to firm-level log wage. To compute labor share, I follow Keller & Yeaple (2009) use firm's sale minus staff expense and operating income (OIBDP) to construct material expense. I then compute the value-added labor share as staff expense divided by sales minus material expense. I keep unique records for each firm and assign a firm to a unique 2-digit industry. When a firm reports both an Industry Format and a Financial Format, I keep the Industry Format; and I exclude firms that do not report an industry code. I eliminate firms with reported cost-of-goods to sales, staff expense to sales ratio, and value-added labor share in the top and bottom 1 percent, where the percentiles are computed for each year separately. The above variables are essential ingredients to measure markups and, also, measure the effect of markup on labor share and wage.

Estimation. De Loecker & Warzynski (2012) proposed a useful way to estimate markups using information from the firm's financial statements. The method does not require any assumptions on market structure or the functional form of the demand that firms face. Instead, this production function approach requires a detailed treatment of the production function to estimate markups by the cost-minimization structure approach.

Following closely De Loecker *et al.* (2020) setup, consider an economy with N firms, indexed by $i = 1, ..., N$. Firms are heterogeneous in their productivity, Z_{it}. Each firm i uses K_{it} units of capital and bundle of variable inputs (including labor,intermediate inputs, materials,...), $\mathbf{V} = (V^1, ..., V^J)$, to produce Q_{it}. Given the production function, firm i minimizes the contemporaneous cost of production in the period t:

$$Q_{it} = Q_{it}(Z_{it}, \mathbf{V}_{it}, K_{it}) \tag{A.1}$$

The key assumption is that within each period, variable inputs adjust freely, whereas capital is subject to adjustment costs and other frictions. The Lagrangian objective function associated with the firm's cost minimization:

$$\mathcal{L}(V_{it}, K_{it}, \lambda_{it}) = P_{it}^V V_{it} + (r_{it} + \delta)K_{it} + F_{it} - \lambda_{it}(Q_{it}(\cdot) - \overline{Q}) \tag{A.2}$$

where P_{it}^V is the price of the variable input, $r_{it} + \delta$ is the user cost of capital. F_{it} is the fixed cost, $Q(\cdot)$ is the technology specified in equation (1), $\overline{Q}$ is a scalar and λ is the Lagrange multiplier. I assume that these input prices are given to the firm. The output elasticity of variable input is given by equation:

$$\theta_{it}^{V} = \frac{1}{\lambda_{it}} \frac{P_{it}^{V} V_{it}}{Q_{it}} \tag{A.3}$$

The estimation procedure is so-called production approach, as in De Loecker & Warzynski (2012) and De Loecker *et al.* (2020). I estimate an approximation of the production function in Equation (A.1) using a sector-year-specific Cobb-Douglas production function, with a variable input bundle and capital as inputs. For a given sector-year st, I consider the production function:

$$q_{it} = \theta_{st}^{V} v_{it} + \theta_{st}^{K} k_{it} + z_{it} + \varepsilon_{it} \tag{A.4}$$

where the lower case letters stand for the log form of the variable represented by the corresponding capital letter. To control for unobserved productivity shocks, which cause simultaneity and selection bias [33], I follow De Loecker & Warzynski (2012) and De Loecker *et al.* (2020) to use control function approach. The unobserved productivity z_{it} is given by a function of the firm's inputs and a control variable, d_{it}, such that $z_{it} = h_{st}(d_{it}, k_{it})$. Following Levinsohn & Petrin (2003) and Ackerberg *et al.* (2015), I use variable input as static control variables to proxy for productivity and assume that the demand for variable input is strictly monotone in z_{it}. Also, as discussed in De Loecker *et al.* (2020), under constant returns to scale assumption, I can identify output elasticities using data on sales and expenditures as long as one can control for markups. I use 4-digit industry market share and productivity as determinants of markups. Therefore, the control variable $d_{it} = (v_{it}, ms_{it})$.

The production function estimation relies on a two-stage approach. In the first stage, the measurement error and unanticipated shocks to output are purged using a non-parametric projection of output on the inputs and the control variable:

$$q_{it} = \phi_t(v_{it}, k_{it}, d_{it}) + \varepsilon_{it} \tag{A.5}$$

where $\phi_t = \theta_{st}^{V} v_{it} + \theta_{st}^{K} k_{it} + h_{st}(d_{it}, k_{it})$. The productivity process is given with law of motion, $z_{it} = g(z_{it}) + \xi_{it}$, given θ_{st}^{V}, I can obtain productivity from $\phi_{it} - \theta_{st}^{V} v_{it} - \theta_{st}^{K} k_{it}$.

In the second stage, I estimate the production function parameters using GMM moment conditions of the following form to obtain the industry-year-specific output elasticity:

$$\mathbb{E}(\xi_{it}(\theta_{st}^{V}) v_{it-1}) = 0 \tag{A.6}$$

This approach identifies the output elasticity of a variable input under the assumption that the variable input will immediately respond to productivity shocks, but capital is subjected to adjustment cost. The persistence in productivity assumption allows variable inputs to be serially correlated.

Wooldridge (2009) proposes to address the unidentified problems in first stage by replacing the two-step estimation procedure with a generalized method of moments (GMM) setup. It suggests estimating both first stage and second stage simultaneously. The estimation proce-

[33] Observed inputs (e.g., labor, capital) may be correlated with unobserved inputs or productivity shocks (e.g., managerial ability, quality of land, materials, capacity utilization). This correlation introduces simultaneity biases. On the other hand, firms exit from the sample is not exogenous and it is correlated with firm size. Smaller firms are more likely to exit than larger firms. Endogenous exit introduces selection-biases.

dure is following: 1) Estimate the two-stage production approach with Ackerberg *et al.* (2015) correction to extract initial values for second step, 2) Estimate the system GMM as proposed in Wooldridge (2009). Furthermore, through out the estimation process, I impose constant return to scale so that it gives the identification of output elasticities using data on sales and expenditure and the estimation setup is line with my model setup. The Wooldridge (2009) methodology regains the efficiency and computes the standard error, which it is hard to achieve in traditional two-stage approach. On the other hand, Flynn *et al.* (2019) points out that applying constant returns to scale reduces the skewness in the markup distribution.

Following De Loecker *et al.* (2020), firm-level markup is defined as output elasticity over its sales to variable cost ratio:

$$\mu_{it} = \theta_s^V \frac{P_{it}Q_{it}}{P_{it}^V V_{it}} \tag{A.7}$$

Since I can directly observe the revenue share of the variable input, $\frac{P_{it}Q_{it}}{P_{it}^V V_{it}}$, the empirical challenge is to recover the output elasticities. Bond *et al.* (2020) points out that using the revenue to estimate the production function does not identify the markup. However, the concern can be addressed by imposing constant returns to scale. Flynn *et al.* (2019) also points out that the markup estimated by De Loecker & Warzynski (2012) approach does not identify it and generate huge variation in markup. Therefore, to obtain the output elasticity of variable input, I impose constant returns to scale because it identifies output elasticities using data on sales and expenditure.

The estimated markup is computed as the output elasticity of variable input. The key assumption is that firms can freely adjust variable input within one period. With detailed data that includes material input and employment, the traditional approach is to assume material expenses are frictionless, as in Levinsohn & Petrin (2003). However, lack of information in material expense pose a critical challenge when using Compustat to estimate production function. To address the issue, I follow the assumption that De Loecker *et al.* (2020) impose on the freely adjusted variable input but further assume that variable input is comprised of constant returns to scale combination of material and labor as follows

$$V_{it} = M_{it}^{\theta^M} L_{it}^{\theta^L}, \ \ \theta_M + \theta_L = 1 \tag{A.8}$$

Departing from frictionless material input, I allow both input markets, material, and labor, to be noncompetitive. I compute labor-variable input share and material-variable input share by Hall (1988) methodology to obtain industry-wide labor cost-share and material cost share. I assume that the shares vary across 4-digit industry and year. I then construct the firm-level material markup and labor markup by the first-order condition

$$\mu_M = \theta^M \theta^V \frac{P_{it}Q_{it}}{P_{it}^M M_{it}} = \frac{\mu_{it}}{\mu_M^d} \qquad \mu_L = \theta^L \theta^V \frac{P_{it}Q_{it}}{w_{it}L_{it}} = \frac{\mu_{it}}{\mu_L^d} \tag{A.9}$$

where $\mu_M^d = \frac{\eta_M}{1+\eta_M}$ and $\mu_L^d = \frac{\eta_L}{1+\eta_L}$ are the markdowns in the material market and labor market respectively. θ^M and θ^L represents the share of material and labor in variable inputs given industry and year. Assuming $\mu_M^d \theta^M + \mu_L^d \theta^L = 1$, I can get

$$\mu_M^d = \frac{1}{\theta^M} \frac{P_{it}^M M_{it}}{P_{it}^V V_{it}} \qquad \mu_L^d = \frac{1}{\theta^L} \frac{w_{it} L_{it}}{P_{it}^V V_{it}} \tag{A.10}$$

In the competitive input market, the input share of variable input should equal the elasticity of variable input and results in $\mu_M^d = \mu_L^d = 1$. Therefore, the markup estimations are identical no matter which variable input is used to construct markup, $\mu = \mu_M = \mu_L$. If input shares deviate from elasticity, the firms would have friction in input markets. I will focus on labor markdown, μ_L^d, only to understand the effects of product and labor market power on labor share and wage.

B Derivation

Wage Distribution. Recall the worker problem:

$$rE(w) = w + \lambda^e \int_{w}^{w_{max}} [E(z) - E(w)]\, dF(z) - \delta(E(w) - U) \tag{B.1}$$

$$rU = b + \lambda^u \int_{w^*}^{w_{max}} [E(z) - U]\, dF(z) \tag{B.2}$$

Evaluating the asset value of employment at w^* and setting it equal to the asset value of unemployment yields:

$$\begin{aligned} w^* &= b + (\lambda^u - \lambda^e) \int_{w^*}^{w_{max}} [E(z) - E(w)] dF(z) \\ &= b + (\lambda^u - \lambda^e) \int_{w^*}^{w_{max}} \frac{1 - F(z)}{r + \delta + \lambda^e[1 - F(w)]} dz \end{aligned} \tag{B.3}$$

We now derive an implicit solution for the wage distribution $G(w)$. A stationary distribution of employment over wages implies:

$$(1 - u)G(w)[\delta + \lambda[1 - F(w)]] = u\lambda_u[F(w) - F(w^*)] \tag{B.4}$$

Evaluating (B.4) at w_{max} yields

$$\frac{u}{1-u} = \frac{\delta}{\lambda_u[1 - F(w^*)]}$$

Substituting into (B.4) gives the solution for $G(w)$:

$$G(w) = \frac{F(w) - F(w^*)}{1 - F(w^*)} \frac{\delta}{\delta + \lambda[1 - F(w)]} \tag{B.5}$$

Finally, we can derive a solution for the mean wage:

$$\bar{w} = \int_{w^*}^{w_{max}} w dG(z) \tag{B.6}$$

Integration by parts yields

$$\begin{aligned} \bar{w} &= w_{max} - \int_{w^*}^{w_{max}} G(z) dz \\ &= [w_{max} - w^*] + w^* - \int_{w^*}^{w_{max}} G(z) dz \\ &= w^* + \int_{w^*}^{w_{max}} [1 - G(z)] dz \\ &= w^* + \frac{\delta + \lambda[1 - F(w^*)]}{1 - F(w^*)} \int_{w^*}^{w_{max}} \frac{1 - F(z)}{\delta + \lambda[1 - F(z)]} dz \end{aligned}$$

Solution. I now outline how I solve the model. Since the production technology is constant-returns-to-scale, I can solve the individual firm problem and evaluate the aggregation results. In particular, first, given a distribution $N(z)$ of the number of firms, I solve for the firm wage $w(z)$ that maximize firm profits. Second, given these choices, we solve for all aggregate quantities and normalize aggregate price $P = 1$ without lose of generality. The profit function in (19) can be rewritten as

$$\pi(z) = \max_{w \geq 0} \left[p(z)z - w(z) \right] l(w) \tag{B.7}$$

subject to labor supply (8) and demand curve (12). The firms choose wage w to post to maximize their profit by solving the first order condition

$$\frac{p(z)z}{\mu(z)} = l\frac{dw}{dl} + w \tag{B.8}$$

The assumption that all workers face the same job arrival rate $\lambda_u = \lambda_e = \lambda$ implies that $w^* = b$ from equation (4). Using the elasticity of labor supply from equation (9) and replacing it to the first term of right-hand side of equation (B.7), I obtain

$$\left[\frac{p(z)z}{\mu(z)} - w \right] \left[\frac{2kF'(w)}{1 + k(1 - F(w))} \right] = 1$$

Focusing on size-invariant equilibria in which wages are increasing in firm productivity, $F(w(z)) = P(z)$. This implies $F'(w(z))w'(z) = P'(z)$ and it follows that the optimal wage relationship between wage and productivity, $w(z)$, solves the ordinary differential equation

$$w'(z) = \left[\frac{p(z)z}{\mu(z)} - w \right] \left[\frac{2kP'(w)}{1 + k(1 - P(w))} \right] \tag{B.9}$$

C Tables

Table C.1 Targeted Moment and Untargeted Moment

			1980	1990	2000	2010	2018
Targeted Moment							
$\mathcal{M}$	cost-weighted	Data	1.18	1.23	1.24	1.26	1.33
	aggregate markup	Model	1.18	1.23	1.24	1.26	1.33
$std(\mathcal{M})$	markup standard deviation		0.23	0.26	0.30	0.33	0.39
			0.23	0.26	0.30	0.33	0.39
$std(\frac{py}{Y})$	relative sales standard deviation		0.87	1.07	1.16	1.06	1.02
	(6-digits NAICS)		0.87	1.07	1.16	1.06	1.02
$\mathcal{M}_d$	sales-weighted		0.99	0.99	0.91	0.79	0.85
	aggregate markdown		0.99	0.99	0.91	0.79	0.85
ρ	replacement rate		0.38	0.37	0.33	0.34	0.33
			0.38	0.37	0.33	0.34	0.33
Untargeted Moment							
W	employment-weighted		3.96	4.01	3.85	3.87	3.67
	aggregate log wage		3.96	3.98	3.96	3.74	3.67
$sd(W)$	log wage standard deviation		0.35	0.40	0.64	0.63	0.63
			0.31	0.32	0.36	0.43	0.41
LS	sales-weighted		0.57	0.56	0.45	0.42	0.44
	aggregate labor share		0.56	0.53	0.49	0.42	0.42

Table C.2 Counterfactual

		1990	2000	2010	2018
		percentage change (%)			
Y	aggregate output	+8.79	+11.54	+2.96	+8.65
	markup	-0.65	-0.84	-1.68	-2.77
	markdown	+1.84	-1.02	-10.62	-5.75
	productivity	+7.60	+13.40	+15.25	+17.17
$\tilde{L}$	production labor	+3.68	-1.91	-17.84	-10.79
	markup	0	0	0	0
	markdown	+3.68	-1.91	-17.84	-10.79
	productivity	0	0	0	0
WL	aggregate labor income	+4.20	-1.87	-22.39	-17.35
	markup	-2.64	-3.52	-5.72	-8.67
	markdown	+3.86	-2.83	-20.39	-13.50
	productivity	+2.99	+4.48	+3.73	+4.81
Z	aggregate productivity	+6.21	+12.98	+17.37	+17.24
	markup	-0.65	-0.84	-1.68	-2.77
	markdown	-0.56	+0.25	+1.66	+1.48
	productivity	+7.42	+13.57	+17.39	+18.53
N	number of firm	+13.81	-28.45	+41.74	+51.31
	markup	+4.63	+6.26	+9.68	+14.56
	markdown	-1.15	+2.55	+6.38	+8.92
	productivity	+10.33	+19.64	+25.69	+27.83
M	entry	+7.37	+33.30	+6.97	+5.63
	markup	+4.63	+6.26	+9.68	+14.56
	markdown	-7.00	+6.66	-22.09	-28.35
	productivity	+9.74	+20.38	+19.39	+19.43

Table C.3 Counterfactual Labor Share Decomposition

		80-90	90-00	00-10	10-18
		average cumulative change (%)			
ΔLS	change in labor share	-0.64	-4.36	-10.44	-13.49
	markup	-0.49	-1.34	-1.95	-2.99
	markdown	+0.78	+0.41	-3.55	-5.24
	productivity	-0.94	-3.43	-4.94	-5.26
$\Delta \bar{LS}$	change in within	-1.11	-3.94	-7.60	-10.72
	markup	-0.45	-1.27	-1.94	-3.15
	markdown	+0.06	-0.10	-1.86	-3.40
	productivity	-0.72	-2.58	-3.79	-4.16
$\Delta cov(ls, \mathcal{SW})$	change in between	+0.46	-0.42	-2.85	-2.78
	markup	-0.04	-0.07	+0	+0.16
	markdown	+0.72	+0.51	-1.69	-1.84
	productivity	-0.22	-0.86	-1.16	-1.10