Monopsony Power and Creative Destruction

January 2026

Isabella Maassen* Filip Mellgren† Jonas Overhage‡

Abstract

This study examines the effects of monopsony on macroeconomic productivity and growth using an endogenous growth model with heterogeneous firms facing upward-sloping labor supply curves. We find that this framework can rationalize the prevalence of unproductive yet innovating firms that would otherwise be crowded out by more productive competitors. In addition, we show that monopsony is an important determinant of product market power. Calibration to U.S. data confirms previous findings that monopsony decreases static productivity. However, it also leads to higher growth. We estimate that a 1% narrowing of the markdown increases the present value of output by about 1.08%. (JEL O31, O40, J42)

I. Introduction

Monopsonistic labor markets decrease aggregate productivity by reallocating labor from the most productive firms to less productive ones. At the same time, monopsony provides a source of economic profits. As these profits incentivize firms to innovate, a trade-off between aggregate productivity and economic growth arises. While product market power has been studied extensively in the endogenous growth literature, other sources of economic profits have received less attention. Meanwhile, empirical evidence for monopsonistic labor markets is mounting, making monopsony a topic for regulation.[1] To examine how monopsony affects the interplay between growth and productivity, this paper develops a tractable endogenous growth model featuring monopsonistic labor markets.

To this end, we introduce monopsonistic labor markets to a Schumpeterian model of

*IIES, Stockholm University, isabella.maassen@iies.su.se

†IIES, Stockholm University, filip.mellgren@iies.su.se

‡IIES, Stockholm University, jonas.overhage@iies.su.se

We thank Timo Boppart, Mitchell Downey, Per Krusell, Simon Mongey, Michael Peters, Florian Trouvain, Joshua Weiss, Horng Wong, and seminar participants at IIES for helpful comments and suggestions.

[1] For example, the 2023 Merger Guidelines (U.S. Department of Justice and Federal Trade Commission [FTC], 2023), considers monopsony as a source of harm when evaluating mergers.

economic growth. This feature implies that firms face upward-sloping labor supply curves. In the model, firms produce a variety of goods with product-specific quality. Firms choose research effort to improve upon the quality of competitors' products, and compete under Bertrand competition for demand in each product line. In equilibrium, demand for any given product is fulfilled by the highest quality producer at a price depending on competitors' marginal costs, and a markup due to its quality advantage. As the degree of monopsony influences the size penalty of growing large, it is important to capture the effect on differently sized firms. We therefore introduce one key heterogeneity: firms can have high or low productivity in production. This feature allows us to examine rich implications of labor market power for the distribution of firm size, marginal costs, and markups.

In turn, the distributions of firm size and markups determine growth and productivity. Productivity improves with labor market competitiveness as process efficient firms are able to set wages that crowd out firms with a low process efficiency. In the case of perfectly competitive labor markets, the only active producer type has the highest process efficiency. Meanwhile, research costs are calibrated to be convex, implying that research is more effective at a small scale. As monopsony enables the survival of many firms that choose to stay small, it stimulates the amount of aggregate innovation. The interaction between monopsony and product market power provides a second channel that stimulates innovation; monopsony enables firms with a high marginal cost to stay active, which results in a higher expected markup following innovation for any given firm type.

To quantify these insights, we calibrate the model to match U.S. data, with the firm-level labor supply elasticity to match micro-level evidence. We then compare the fully calibrated model to a restricted version of the model without monopsony. Targeted moments for remaining parameters include the unemployment rate, key labor market elasticities, markups, the growth rate, and the revenue share of the top 10% largest firms. We then validate the monopsonistic model against the competitive benchmark using untargeted data moments, such as the research and profit shares, as well as relative process efficiencies and wages. With regards to firm-level productivity – or process efficiency as we designate it in this paper – the monopsonistic model is able to more closely match the disparity observed in the data. This occurs because monopsonistic producers with a low productivity can keep marginal costs relatively low by staying small, which is not possible when all firms pay the same market wage, as in the perfectly competitive case.

We use the calibrated model to investigate how the degree of monopsony affects the present value of aggregate output. We separate effects into four components: growth, average process efficiency, markup dispersion, and employment. We find that the average process efficiency and the rate of creative destruction are the biggest drivers of the overall effects. A one percent narrowing of markdowns leads to a 1.97% increase in average process efficiency, and a 1.16% reduction in the present value of future output due to lower growth. Decreasing markup dispersion improves output by 0.26%, and aggregate employment increases modestly by about 0.01%.

We connect to the policy debate by extending the model to feature progressive income taxes. We first show that lower income tax progressivity increases the firm-level labor supply elasticity. We then apply the model to a historical tax reform and estimate that the tax cuts in the 1980s increased the firm-level labor supply elasticity by about 2.3%, contributing to modest improvement in U.S.output. Next, we perform a hypothetical policy exercise where we fix the level of tax revenue at a base level. We find that, similar to directly setting the elasticity, a low-progressivity regime with high market concentration maximizes the net present value of output at the cost of its long-run growth rate.

Altogether, we obtain a rich framework that allows us to understand how the firm-level labor supply elasticity shapes aggregate output through its effect on entry, the firm size distribution, misallocation, and employment. Furthermore, we speak to how these elasticities can be affected by tax policy. Our framework can also be used to study questions related to long-run developments in labor markets.

The paper proceeds as follows. First, the relevant literature is discussed in Section A. Then, the model is introduced in Section 2. Section 3 shows how the equilibrium can be characterized. Next, Section 4 discusses the analytical results before turning to a quantification of the model. The extension with a progressive tax schedule is introduced and discussed in Section 5, before we conclude in Section 6.

A. Related literature

Our paper first and foremost relates to the literature on endogenous growth, building on seminal work such as Aghion and Howitt (1992) and Klette and Kortum (2004). In particular, we base our model on Aghion et al. (2023) and adapt it in several ways, with the main difference being the addition of monopsonistic labor markets. To our knowledge, ours is the first paper to incorporate monopsony power in the labor market for production workers in this class of models featuring endogenous growth from creative destruction.

Closely related work includes Peters (2020) and De Ridder (2024). Peters (2020) develops a growth model in which markups arise endogenously following risky own-innovations that improve the quality of products already under control of the producer. De Ridder (2024) develops an alternative model in which firms adopt intangible technology that lowers the marginal cost of production. The main similarity between our paper and these two contributions is that low marginal cost firms are able to charge higher markups which provide incentives to innovate. Another parallel is that as the economy becomes increasingly dominated by a few large firms, growth falls following decreasing returns to research. In contrast to these contributions, in our model marginal costs are not determined by previous innovations or investment, but increase directly with output as a result of monopsony. In addition, markups arise endogenously in our model depending on the equilibrium firm size distribution.

Other recent contributions to the literature on creative destruction and market power include Cavenaile et al. (2019), studying strategic investments into productivity growth under oligopolistic competition; Peters and Walsh (2021), which considers how population growth affects productivity growth and finds that decreasing population growth reduces competition and increases markups; Liu et al. (2022), study how a low interest rate environment can increase market concentration and its effect on productivity growth; Focusing on business dynamism, Akcigit and Ates (2023) document a number of trends such as increased markups and concentration, and a fall in the labor share, and build a general equilibrium model featuring creative destruction to jointly explain these trends. In particular, they emphasize the role of a decrease in the intensity of knowledge diffusion between frontier and laggard firms. Finally, the work by Weiss (2023) analyzes large firm innovation incentives and finds that increasing large firm profitability helps explain the recent growth slowdown in the U.S. The main contribution of this paper relative to these papers is the focus on monopsonistic labor markets and the interactive role that plays for determining product market competition.

Studies on monopsony that consider economic growth include work by Garibaldi and Turri (2024) and Fernández-Villaverde et al. (2025). The former study monopsony in a neoclassical growth model and corroborate the finding that monopsony lowers the level of output while increasing growth. The latter investigate how monopsony in the market for researchers affects growth when incumbent producers have a strategic capacity to protect their product lines from entrants. Similarly, Lehr (2024) examines monopsony power in the market for inventory and finds that monopsony power in this market reduces US economic growth and welfare. In relation to these contributions, our paper instead focuses on the effects of monopsony in the market for production workers.

On the macroeconomic consequences of market power, several studies point to the harm on static productivity. Berger et al. (2022) study oligopsony in a general equilibrium model and find that aggregate output is about 20.9% lower due to labor market power compared to a competitive benchmark. Baqaee and Farhi (2019), and Edmond et al. (2023) quantify large costs to productivity arising from markups due to the level, dispersion, and effects on entry. Bachmann et al. (2024) study variation in labor market power between East and West Germany, and find that variation in labor market power can explain about 40% of the difference in productivity between the two regions. Relative to these papers, our paper contributes by developing a structural model that can be used to study product innovation and growth in addition to static output. Due to this feature, the model is also able to quantify the trade-off between static productivity and dynamic efficiency that arises.

The empirical evidence for monopsonistic labor markets has seen a recent surge. Notably, Sokolova and Sorensen (2021) synthesize a large body of micro-level studies which estimates the labor supply elasticity firms face when posting wages and arrive at a point estimate around 7.1 implying wage markdowns arising from monopsony around 88%. Another comprehensive study that attempts to identify the firm-level labor supply elasticity from the worker separation responses to firm wage policies is Bassier et al. (2022), which find a firm-level labor supply elasticity around 4.2 and wage markdowns around 81%.

A different strand of the literature finds empirical evidence for labor market power in matched employer-employee data, e.g. Lamadon et al. (2022); or using a production function approach, Chen et al. (2022), Kirov and Traina (2023) for the case of U.S. manufacturing, and Estefan et al. (2024) in the case of Mexican manufacturing. A recurring theme is substantial and widespread wage markdowns.

In modeling the labor market, we build on Card et al. (2018), where the source of wage setting power is workers' idiosyncratic preferences over workplaces. These idiosyncratic preferences give rise to an upward sloping labor supply curve for each firm, as workers weigh wages and non-wage preferences. This approach to modeling monopsony is what Manning (2021) classifies as "new classical monopsony." A distinctive feature of this framework is wage posting; employers post wages and hire any worker willing to accept their posted wage. There is no scope for wage bargaining. Consequently, each employer sets a firm-wide wage. Moreover, workers observe all wage offers and don't have to draw a subset of offers as is common in search models. For a thorough review of wage posting and monopsony Kline (2025) provides a comprehensive overview. Finally, this approach to modeling labor markets also contrasts other models where firms' wage setting power comes from local labor market power, that is a firms' hiring share in the local labor market with a (small) finite amount of employers, as in Azkarate-Askasua and Zerecero (2024).

One feature of our model is that equilibrium wage inequality between workers stems from firm-wide wage premia, which is consistent with empirical studies such as Abowd et al. (1999), Bonhomme et al. (2019), Bonhomme et al. (2023) or Wong (2023). We additionally speak to the literature on developments in the labor share, such as work by Elsby et al. (2013), Karabarbounis and Neiman (2014) or Rodriguez and Jayadev (2013). It may seem intuitive that higher monopsony power would lead to a decrease in the labor share. However, our model of creative destruction uncovers another channel: while monopsony increases the profit share for any given firm, it also shifts economic activity towards firms with a relatively low profit share. Accounting for both channels makes the relation between monopsony and the labor share ex-ante ambiguous.

II. Model

Like standard models of creative destruction, our model features monopolistic firms with production of intermediate goods along a quality ladder. Intermediate goods are bundled by a competitive final goods producer. Households value final good consumption and have idiosyncratic preferences over different employers. Thus, households make a discrete choice regarding their workplace and employment status. The economy grows as a result of firms' innovation efforts that lead to an increasing quality level of goods.

A. *Final good production*

There is a competitive final goods producer that aggregates differentiated intermediate goods from a unit interval according to a Cobb-Douglas aggregator:

$$Y_t = \exp \int_0^1 \ln(q_{it} y_{it}) di, \quad i \in [0, 1], \tag{1}$$

where Y_t is final output, q_{it} is the quality level of good i, and y_{it} is the quantity of that good. This set-up implies that demand for each differentiated good follows:

$$p_{it} y_{it} = P_t Y_t, \quad P_t \equiv \exp \int_0^1 \ln(p_{it}/q_{it}) di, \tag{2}$$

where we normalize $P_t \equiv 1$. For a detailed derivation of intermediate good demand, refer to Appendix A.1. We further introduce a quality index $Q_t \equiv \exp \int_0^1 \ln(q_{it}) di$, such that aggregate output can be interpreted as a combination of quality level and physical output: $Y_t = Q_t \exp \int_0^1 \ln(y_{it}) di$.

This demand specification makes intermediate goods in the same product line i produced by different firms perfect substitutes and the final goods producer purchases from the firm j with the lowest quality-adjusted price, i.e. $p_{it}/q_{it} = \min_{j \in \mathcal{J}} p_{ijt}/q_{ijt}$. To break ties between intermediate goods producers posting equal quality-adjusted prices, we assume that the good with the higher quality is preferred.

For technical details on the good demand, refer to Appendix A.2. Essentially, intermediate goods producers compete in a Bertrand manner within each product market, and product demand for good i facing firm j is formally expressed as:

$$y_i(p_{ijt}, q_{ijt}, Y_t) = \begin{cases} \frac{Y_t}{p_{ijt}} & \text{if } \frac{p_{ijt}}{q_{ijt}} < \frac{p_{ij't}}{q_{ij't}}, \forall j' \in \mathcal{J} \setminus j \\ \frac{Y_t}{p_{ijt}} & \text{if } \frac{p_{ijt}}{q_{ijt}} \leq \frac{p_{ij't}}{q_{ij't}}, \forall j' \in \mathcal{J} \setminus j \wedge q_{ijt} > q_{ikt}, \forall k : \frac{p_{ijt}}{q_{ijt}} = \frac{p_{ikt}}{q_{ikt}} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

B. Households

A mass $\mathcal{L}$ of households derive utility from private consumption and make a discrete choice over workplaces and home production in each period. They can choose to seek employment ($g = e$) in a company $j \in \mathcal{J} = \{1, ..., J\}$, or engage in home production ($g = u$). Households, indexed by o, have preferences for consumption as well as working at different firms and home production,

$$u_{ojt} = \beta \ln C_{ojt} + \xi_{ogt} + (1 - \sigma)\epsilon_{ojt}, \tag{4}$$

where C_{ojt} is household consumption. Households do not save, but fully consume the wage earned at firm j, which implies $C_{ojt} = W_{ojt}$. In our model, wages will differ only across firms j, not across households working within a single firm. We thus drop the o subscript for the wage, and can state household utility as:

$$u_{ojt} = \beta \ln(W_{jt}) + \xi_{ogt} + (1 - \sigma)\epsilon_{ojt}, \tag{5}$$

where ϵ_{ojt} is independently and identically extreme value type 1 distributed. Similarly, $\xi_{ogt} + (1 - \sigma)\epsilon_{ojt}$ is also i.i.d. extreme value type 1 distributed. If the household works at a firm, they earn the firm j specific wage W_j. If they engage in home production, they instead receive ωY_t. The presence of the outside option helps identify the wage level.

Households compare all options available to them and choose to work where they receive the highest utility. The formulation thus captures, in addition to wages, individual preferences over working at any given firm (ϵ_{ojt}) and being employed at all (ξ_{ogt}), which does not depend on the workplace.

The parameter β captures how sensitive a household's utility is to consumption. In the limiting case where $\beta \to \infty$, households care only about consumption and choose their workplace based on where they earn the highest wage. This case represents perfect competition and the economy's labor market can be said to be perfectly competitive.

The parameter σ determines the weight $1-\sigma$ a household places on idiosyncratic workplace preferences. In the limiting case where $\sigma \to 1$, households choose a workplace based on the wage alone, similar to the case $\beta \to \infty$ discussed above. Labor markets are competitive in either case, but with different implications for unemployment as discussed below.

Having the ability to choose the outside option of engaging in home production serves as an anchor for the wage level. Moreover, one intuition for this set-up is that a lower wage level at their preferred firm means workers become more likely to opt for home production instead of employment, where the parameter σ measures the sensitivity of this trade-off with respect to idiosyncratic preferences over firms.

The nested logit set-up follows the model outlined in McFadden (1977). We formulate the model with an outside option allowed to grow along a balanced growth path, which ensures a unique equilibrium. The implied labor supply facing firm j is given by:

$$L_j(W_{jt}) = \mathcal{L}\frac{W_{jt}^{\frac{\beta}{1-\sigma}}}{(\sum_{k=1}^{\mathcal{J}} W_{kt}^{\frac{\beta}{1-\sigma}})^{\sigma}(\omega Y)^{\beta} + \sum_{k=1}^{\mathcal{J}} W_{kt}^{\frac{\beta}{1-\sigma}}},$$

For details on solving the nested discrete choice problem, refer to Section A.3 in the Appendix. The labor supply depends on the amount of workers $\mathcal{L}$, and the wage of firm j relative to other firms' wages and the outside option ωY. Labor supply increases in the own wage, and decreases in competitors' wages. The own-wage elasticity of labor supply is $\beta/(1-\sigma)$. We define $D_{e,t} \equiv \sum_{k=1}^{J} W_{kt}^{\frac{\beta}{1-\sigma}}$ and $z_t \equiv \mathcal{L}/\left(D_{e,t}^{\sigma} \cdot (\omega Y)^{\beta} + D_{e,t}\right)$. The labor supply facing firm j is then given by:

$$L_j(W_{jt}) = z_t W_{jt}^{\frac{\beta}{1-\sigma}}, \tag{6}$$

where z_t is an equilibrium object taken as given by firms, assuming that each individual firm considers itself small enough to not have an influence on z_t. Firms hence do not take the effect of their wage on the labor market into account, making the model a model of monopsony as opposed to oligopsony. The ratio $\frac{\beta}{1-\sigma}$ measures the labor supply elasticity that firms face and a lower value corresponds to a greater degree of monopsony.

To build intuition about why both σ and β are introduced, we derive the elasticity of the share of workers choosing home production with respect to the value of the outside option, ωY. Summing up the labor employed at all firms yields the rate u_t of workers choosing the outside option $u_t \equiv 1 - \sum_{j\in\mathcal{J}} L_j/\mathcal{L}$. Applying this, the sensitivity of the home employment rate with respect to the outside option value is given by:

$$\varepsilon_{u,\omega Y} = \frac{\partial u}{\partial(\omega Y)}\frac{(\omega Y)}{u} = \beta(1-u). \tag{7}$$

That is, β controls how strongly the home employment rate responds to changes in the value of the outside option, where the response becomes less sensitive for higher values of β. More details about the derivation of this elasticity can be found in Appendix A.4.

C. Intermediate goods producers

Intermediate goods producers, $j \in \mathcal{J}$, set prices for intermediate goods, and decide how much to invest into research. The firm problem can conceptually be divided into two optimization problems: a static and a dynamic one. Statically, setting prices for intermediate goods determines production quantities and profits within a period. Dynamically, the firm decides how much to invest in research today, which leads to quality innovations in the next period. In the Bertrand Nash equilibrium, the firm that produces the highest quality of an intermediate good, is able to sell it at a marked-up price. In the following, the two parts of the firm problem are described in more detail.

Within a period, firms maximize their static profit by setting their quality-adjusted price for each intermediate profit, p_{ijt}/q_{ijt}, taking as given the current state of product quality in each line, q_{ijt}, as well as the quality adjusted prices of rival firms. As competitors' prices follow from their respective marginal costs and qualities, firm profits depend on the distributions of mc_j, q_j, denoted $\Gamma_{mc,t}$ and $\Gamma_{q,t}$ respectively. Firms produce physical output using a single input referred to as labor using a firm-specific production technology, $y_j = s_j f(l_j)$, which depends on the firm's process efficiency s_j. A firm's process efficiency is the only source of heterogeneity in this model. As described above, intermediate product demand is given by Equation 3. Setting intermediate prices hence determines how much the firm produces, which in turn implies the required labor input. The wage is then set via the labor supply facing the firm, such that the labor input is matched. Formally, the static problem is given as:

$$\Pi_j(Y_t, \{q_{ijt}\}_{i\in[0,1]}, \Gamma_{mc,t}, \Gamma_{q,t}) = \max_{\{p_{ijt}\}_{i\in[0,1]}} \int_0^1 p_{ijt} y_{ijt} di - W_{jt} L_{jt}, \tag{8}$$

$$\text{s.t.} \quad L_{jt} = \frac{1}{s_j} \int_0^1 f^{-1}(y_{ijt}) di, \quad W_{jt} = \left(\frac{L_{jt}}{z_t}\right)^{\frac{1-\sigma}{\beta}} \tag{9}$$

$$\text{Intermediate product demand as in (3).} \tag{10}$$

To impact its future ability to make static profits via $\{q_{ijt}\}_{i\in[0,1]}$, the firm can choose to engage in research. Denote the last firm to innovate upon product line i as $j(i)$ and the firm with the second-highest quality as $j'(i)$. We refer to these firms as "(quality) leader" and "(quality) follower." Innovation is modeled as a $\gamma > 1$ step over the highest existing quality level of a good, i.e., $\gamma = q_{j(i)}/q_{j'(i)}$. Research is undirected in the sense that firms do not decide which product lines to innovate in. However, we assume that firms do not innovate on product lines where they are already the leading quality producer, and that they do not draw the same product line as someone else in the same period. Research costs C^R are a function of the mass of product lines x_{jt} the firm wishes to become the quality leader in. The full firm problem is:

$$V_{jt}(\{q_{ijt}\}_{i\in[0,1]}) = \max_{x_{jt}} \Pi_j\left(Y_t, \{q_{ijt}\}_{i\in[0,1]}, \Gamma_{mc,t}, \Gamma_{q,t}\right) - C^R\left(Y_t, x_{jt}\right) \tag{11}$$

$$+ R_t V_{t+1}(\{q_{ijt+1}\}_{i\in[0,1]}). \tag{12}$$

Where $\Gamma_{x,t}$ is a distribution objects containing the full time t distribution of firms' expansion choice x. Note that $\Gamma_{mc,t}, \Gamma_{q,t}$ are co-determined with Y_t within-period in general equilibrium. On the other hand, $\Gamma_{x,t}$ follows a law of motion:

$$\Gamma_{q,t+1} = H(\Gamma_{q,t}, \Gamma_{x,t}). \tag{13}$$

With each firm owning a mass of product lines, its own quality levels $\{q_{ijt}\}_{i\in[0,1]}$ and the distributions $\Gamma_{q,t}, \Gamma_{x,t}$ are sufficient to make decisions about research, x_{jt}.

Firm entry and ownership. Firms are owned by absentee capitalists with discount factor ρ. This implies that firms' value functions are discounted at rate $R_t = \rho/g_t$. For details on this, refer to Appendix A.5.

When a firm enters, it pays a cost ζY_t and draws its type b. It additionaly becomes the quality leader in a mass of lines equal to n_b^*, mirroring the average firm of its type in the economy. Note that households now draw idiosyncratic preference shocks for working at that firm. The firm starts producing output and can invest in R&D to grow. Firms will hence enter as long as the expected firm value is greater than the entry cost:

$$\zeta Y_t \leq E_t \left[V_{bt}(n_{bt}^*) \right]. \tag{14}$$

With this set-up, entry costs scale with output following Klenow and Li (2025). An effect of this modeling choice is that policies that improve aggregate output by making labor markets more competitive will not see increased entry. We further assume that firms enter at their optimal size upon paying the entry costs. Note that there are no firm-level shocks. This property, combined with the output-scaling of the entry cost implies that there will be no entry on the BGP. In such an equilibrium, the number of firms is thus constant.

D. *Market clearing*

Final output Y_t is used for research expenditure $\tilde{C}_t$ and consumption, which occurs in the form of wage-financed private consumption C_t, and rents E_t. The following identity must hold:

$$Y_t = \tilde{C}_t + C_t + E_t. \tag{15}$$

In this model, there are rents due to non-zero entry costs ζY_t. We can express research costs as:

$$\tilde{C}_t = \sum_j C_t^R(X_t n_{jt}). \tag{16}$$

Private consumption is the sum of all net wages paid:

$$C_t = \sum_{jt} W_{jt} L_{jt}, \tag{17}$$

and rents are the sum over profits in production minus research costs:

$$E_t = \sum_j \left(\Pi_j(Y_t, \{q_{ijt}\}_{i\in[0,1]}, \Gamma_{mc,t}, \Gamma_{q,t}) - C^R(X_t n_{jt}) \right). \tag{18}$$

III. Characterizing the equilibrium

In this section, we show how to solve the model and how to characterize key equilibrium objects, including wages, total factor productivity and aggregate employment. We begin by introducing three assumptions making the analysis tractable. First, we assume that the intermediate goods production is linear in labor, that is $y_j(l_{ijt}) = s_j l_{ijt}$.

Second, we assume that research costs can be approximated by the following function:

$$C^R(Y_t, x_{jt}) = \psi Y_t (x_{jt})^{\phi}, \tag{19}$$

where ϕ governs the returns to scale of R&D investment. In our calibration, we will find $\phi > 1$ matches the firm size distribution and relative markups, i.e. research costs are convex and smaller firms are endogenously more productive at doing research.

Third, we assume that there are two types of intermediate goods producers that differ in their process efficiency, $s_j \in \{s_L, s_H\}$, $s_L < s_H$. Allowing for different firm types enables us to showcase how imperfect competition in the labor market affects aggregate output by enabling less process efficient firms to survive. We denote the share of type H firms by α, and the share of product lines held by them as h_t at time t, or h^* along a balanced growth path. Note that α is a parameter determining the share of firms that belong to either type, and that h^* measures market concentration in equilibrium. We will later consider equilibrium solutions where $\alpha < h^* \leq 1$, i.e. the more process efficient firms are larger than the less process efficient firms.

The focus in the remainder of this section is on the economy's steady state. Although the model admits firm dynamics and short run transitions between the intermediate producers' state variables, we are going to focus on a steady state in which the economy grows along a balanced growth path (BGP) defined as:

Definition 1 (Balanced growth path equilibrium)**.** *A balanced growth path equilibrium is an equilibrium in which all variables and prices grow at constant rates, and the number of active firms is constant.*

We will proceed by first discussing implications of this definition for the within-period (static) optimization, before solving for the dynamic optimization and BGP equilibrium conditions.

A. Within-period optimization

In the Bertrand equilibrium, intermediate goods producers have the ability to earn positive profits in the mass n_{jt} of lines where they are currently product leaders by setting the price in each line where they are quality leaders equal to the quality improvement times the marginal cost, mc_j, of the quality follower: $p_{it} = \gamma \cdot mc_{j(i)t}$[2]. Quality followers set prices equal to their own quality-adjusted marginal cost. In Appendix A.2 we show that this is a Nash equilibrium. Pricing according to the scheme described above attracts demand for goods in each line equal to $y_{it} = PY/p_{it}$, which means the intermediate producer's revenue equals the mass of lines where the firm is a quality leader multiplied by the size of the economy $n_{jt}Y_t$. The line-level price depends on the probabilities that the follower is of either firm-type and will in expectation equal $\mathbb{E}(p_i) = \gamma \sum_{j \in \mathcal{J}} n_j mc_j$. Assuming there is no uncertainty over the future state of the economy, i.e. the path of Y_t and R_t is known, firms will only engage in costly research if there is something to gain from being a quality leader in additional lines. In other words, firms do not invest into gaining a quality advantage if they do not plan to use that advantage to generate (static) profits. This implies that firms will produce in all lines in which they are quality leaders, given that they behaved optimally in their dynamic optimization.

Definition 2 (Static Equilibrium). *Taking quality levels and prices as given, firms maximize profits by choosing firm-level employment and how to allocate it and households maximize utility by selecting where to work.*

In all lines i, the quality leader $j(i)$ sets the quality-adjusted price equal to his follower's quality-adjusted marginal costs, that is $p_{ij(i)t} = \gamma mc_{ij'(i)t}$, and fulfills the implied product demand y_{it}. The firm size in terms of the number of varieties produced n_{jt} is hence given as:

$$n_{jt} = \int_0^1 \mathbf{1}(q_{ijt} > q_{ikt}, \forall k \in \mathcal{J} \setminus j) di. \tag{20}$$

From here, revenue is simply given as $n_{jt}Y_t$ as revenue per line is Y_t from intermediate goods demand: $p_{ijt}y_{ijt} = Y_t$. The quantity that needs to be produced in each line is then $y_{ijt} = Y_t/(\gamma mc_{j'(i)t})$. Total physical output that the firm needs to produce is hence given as:

$$Y_{jt} = \int_0^1 \mathbf{1}(q_{ijt} > q_{ikt}, \forall k \in \mathcal{J} \setminus j) \cdot Y_t/(\gamma mc_{j'(i)t}) di = n_{jt} \frac{Y_t}{\gamma} m_t^{-1}, \tag{21}$$

[2]See Appendix A.6 for derivations of marginal cost.

where m_t is the harmonic mean of marginal costs:

$$m_t^{-1} \equiv h_t/mc_{h,t} + (1-h_t)/mc_{l,t}. \tag{22}$$

From the firm-level quantity, the flow profit function given $n_{j,t}, m_t$ can now be simplified to:

$$\Pi_j^{BGP}(Y_t, n_{jt}, m_t) = n_{jt}Y_t - L_{jt}W_{jt}, \tag{23}$$

$$\text{s.t. } Y_{jt} \text{ given by Equation 21,} \tag{24}$$

$$L_{jt} = \frac{Y_{jt}}{s_j}, \quad W_{jt} = \left(\frac{L_{jt}}{z_t}\right)^{\frac{1-\sigma}{\beta}}. \tag{25}$$

$$\tag{26}$$

Where z_t is a function of Y_t according to its definition.

B. Dynamic optimization

The dynamic firm problem amounts to choosing n_{jt+1}. Note that other firms will become quality leaders in some of the product lines included in n_{jt} in the next period. To take into account this creative destruction by other firms, we introduce the aggregate variable $X_t = \sum_{j\in\mathcal{J}} x_{jt}$, and the stock n_{jt} of lines that firm j is the quality leader in decreases at that rate.

Along its BGP, growth in this model is driven by creative destruction through quality improvements. As shown below, the level of aggregate quality $Q_t \equiv \exp\int_0^1 \ln(q_i)di$, depends on the quality step size γ and the aggregate rate of creative destruction, X_t:

$$g_Q = \exp\int_0^1 \ln(q_{it+1}/q_{it})di = \exp[X_t\ln(\gamma) + (1-X_t)\ln(1)]di = \gamma^{X_t}, \tag{27}$$

meaning that the rate of creative destruction, X_t, must be constant on the BGP.

Growth in final output, Y_t, is denoted by $g_t \equiv Y_{t+1}/Y_t$ and is entirely driven by quality improvements. Rewriting Equation 1, we can express the growth of final output as:

$$g = \frac{Y_{t+1}}{Y_t} = \frac{Q_{t+1}}{Q_t}\frac{\int_0^1 \ln(y_{jt+1})di}{\int_0^1 \ln(y_{jt})di} = g_Q. \tag{28}$$

Note that a requirement is that average production per intermediate product stays constant on the BGP, which is the case if the producer type distribution h^* is constant. We solve for growth rates of all other variables in Appendix A.8 and summarize our findings below:

$$\begin{aligned} g_z &= g^{-\frac{\beta}{1-\sigma}}, \\ g &= g_Y = g_Q = g_m = g_w = \gamma^X. \end{aligned}$$

These insights on BGP conditions also imply that Γ_q is no longer needed in the profit function, as firms only need to know the share of followers of either type to determine prices, and therefore profits. Recall that the revenue share of H-type followers is h_t, and the share of L-type followers is $1 - h_t$. As this distribution is stable for all firms on BGP, only the index of marginal costs m_t^{-1}, is needed, rather than the whole distribution. On a balanced growth path, intermediate producers maximize the following:

$$V_{jt}(n_{jt}) = \max_{x_{jt}} \Pi_j^{BGP}(Y_t, n_{jt}, m_t) - C^R(Y_t, x_{jt}) + R_t V_{jt+1}(n_{jt+1}) \tag{29}$$

$$\text{s.t. } n_{jt+1} = (1 - X_t) n_{jt} + x_{jt}. \tag{30}$$

To continue solving for the balanced growth path equilibrium, it will be helpful to solve for the intermediate goods producers' firm-level markups, which are aggregated up from the markups in product lines. These are equal to the product line price relative to the intermediate good producer's marginal cost, $\mu_{ijt} \equiv p_{ijt}/mc_{jt}$. We can therefore define markups at the firm-level, μ_{jt}, as the quantity-weighted average price over the firm's marginal cost:

$$\mu_{jt} \equiv \frac{\int_0^{n_{jt}} y_{it} p_{it} di}{mc_{jt} \cdot \int_0^{n_{jt}} y_{it} di} = \frac{\gamma m_t}{mc_{jt}}. \tag{31}$$

Markups are constant along a balanced growth path and we denote their BGP values as μ_L^* and μ_H^* respectively.

Recall that m_t^{-1} is defined as a harmonic mean of marginal costs $m_t^{-1} \equiv h_t/mc_{h,t} + (1 - h_t)/mc_{l,t}$. This definition can be rewritten to provide an equation linking markups and the firm size distribution. Being a definition, it holds regardless of whether the economy is on the BGP and we rewrite it as follows[3]:

Lemma 3.1. *The marginal cost index, m_t, can be rewritten in terms of markups and the firm size distribution in the following way:*

$$-\frac{\gamma - \mu_{Lt}}{\gamma - \mu_{Ht}} = \frac{h_t}{1 - h_t}. \tag{E1}$$

For parameters where $\alpha < h^* < 1$, we have the following relation between the markups: $1 < \mu_L^* < \gamma < \mu_H^*$. In the case where $h_t \to 1$, the economy is controlled entirely by the producer type with the highest process efficiency.

With both firms active in equilibrium, the relative wage W_H^*/W_L^* is well defined along the BGP. In addition, due to upward sloping labor supply curves, the relative wage is informative of the relative firm size, and we can use the ratio to obtain another equation which links the three equilibrium outcomes h^*, μ_L^*, μ_H^*.

[3]See Appendix B for proofs of all lemmata.

There are two ways to solve for wages. First, we can use the Bertrand Equilibrium to obtain the firm size which gives a wage expression. In the Bertrand Nash equilibrium, the quality leader chooses to always fulfill demand at the price determined by the quality follower's marginal cost, marked up by the constant factor γ. See Appendix A.2 for details. Combined with the assumption of firm productivity, we obtain total labor input at the firm-level: $L_{jt} = \int_0^1 y_{ijt}/s_j di$. From intermediate goods demand, we know $y_{it} = Y_t/p_{it}$ and from the static problem, we have the line level price given by $p_{ij(i)t} = \gamma mc_{ij'(i)t}$. Firm-level labor can be aggregated according to the following:

$$L_{jt} = \frac{Y_t}{\gamma s_j} \int_0^1 \frac{1}{mc_{j'(i)t}} di = \frac{Y_t n_{jt}}{\gamma s_j} m_t^{-1}. \tag{32}$$

Via the labor supply function, this implies wages are given by:

$$W_{jt} = (L_{jt}/z_t)^{\frac{1-\sigma}{\beta}} = \left(\frac{Y_t n_{jt}}{\gamma s_j m_t z_t} \right)^{\frac{1-\sigma}{\beta}}. \tag{33}$$

The size of a firm is thus determined by the mass of products they span and the marginal costs of competitors. When competitors' marginal costs are high, firms will hire fewer workers due to their ability to charge higher prices and therefore produce less.

Second, firm-level marginal costs are defined as the change in cost as firm-level output increases.[4] Taking the derivative of firm costs with respect to Y_{jt} yields marginal cost:

$$mc_{jt} = \frac{\partial C_{jt}}{\partial Y_{jt}} = \left(\frac{1-\sigma}{\beta} + 1 \right) W_j \frac{1}{s_j}. \tag{34}$$

The marginal cost expression above can be rearranged for a second expression of the firm-level wage:

$$W_{jt} = s_j \cdot mc_{jt} \cdot \frac{\frac{\beta}{1-\sigma}}{1 + \frac{\beta}{1-\sigma}}. \tag{35}$$

This expression shows how the labor supply elasticity facing intermediate goods producers affects the wage markdown. Note that the total markdown is given by $\lambda_j \equiv W_j/mrpl_j$, where $mrpl_j$ is the marginal revenue product of labor. In this model, the marginal revenue product of labor depends on the markup and the labor supply elasticity:

$$\lambda_j = \mu_j^{-1} \frac{\frac{\beta}{1-\sigma}}{1 + \frac{\beta}{1-\sigma}}. \tag{36}$$

Since part of the wedge between the wage and the marginal revenue product of labor comes from the fact that producers can set a markup over marginal cost, μ, on the product market, we define a 'pure' monopsony markdown ν, which is only driven by the fact that

[4] Details on deriving marginal costs are in Appendix A.6.

the labor supply curve is upward sloping:

$$\nu \equiv \mu\lambda = \frac{\frac{\beta}{1-\sigma}}{1+\frac{\beta}{1-\sigma}}. \tag{37}$$

In the following, we refer to ν as the markdown or "net markdown." Combining the the two wage expressions gives the second relation between h^*, μ_L^*, μ_H^*. Using Equation 35 and the identities $\alpha n_H J = h$, $(1-\alpha)n_L J = 1-h$, we establish Lemma 3.2 as shown in Appendix B.

Lemma 3.2.

$$\frac{\alpha}{1-\alpha}\left(\frac{s_H}{s_L}\right)^{\frac{\beta}{1-\sigma}+1} \cdot \left(\frac{\mu_{Lt}}{\mu_{Ht}}\right)^{\frac{\beta}{1-\sigma}} = \frac{h_t}{1-h_t} \tag{E2}$$

Notice that this is the key equation where the degree of monopsony $\frac{\beta}{1-\sigma}$ enters into the determination of (μ_L^*, μ_H^*, h^*).

Next, we restate the dynamic problem of the intermediate producers in Equation 29, and use the objective function's first order condition to arrive at a third equilibrium condition. We rewrite the problem by first substituting the constraint, $x_t = n_{t+1} - (1-X_t)n_t$, and then dividing both sides by Y_t. Moreover, we note that $R_t = \rho/g$[5] and arrive at a simplified value function in detrended terms, where $\tilde{V}_j = V_{jt}/Y_t$. In the following, we drop the time subscripts to emphasize that all variables are detrended:

$$\begin{aligned}\tilde{V}_j(n_j) = &\max_{n_j'} n_j - \left(\frac{n_j}{s_j\gamma m}\right)^{\frac{1-\sigma+\beta}{\beta}}\left(\frac{Y}{z}\right)^{\frac{1-\sigma}{\beta}} - \psi(n_j' - (1-X)n_j)^{\phi} \\ &+ \rho\tilde{V}_j(n_j').\end{aligned} \tag{38}$$

Notably, monopsonistic labor markets affect the firm problem via the shape of the production cost function. This value function implies that that the firm chooses a markup and firm size, which depend on the labor supply parameters σ, β, to balance out the research costs necessary to be paid to be of that size. Rearranging the first order condition of the BGP firm problem in Equation 38 and imposing the BGP condition $n_j' = n_j$ yields an expression for optimal firm size:

Lemma 3.3.

$$n_j^* = \left(\frac{\mu_j^*-1}{\mu_j^*} \cdot \frac{1}{\phi\psi} \cdot \frac{1}{(X^*)^{\phi} + (X^*)^{\phi-1}\cdot\frac{1-\rho}{\rho}}\right)^{\frac{1}{\phi-1}}. \tag{FOC$_j$}$$

Similar to the discussion of the value function above, the first order condition reveals a tight relationship between markups μ_j, the firm size n_j, and the rate of creative destruction

[5] See Appendix A.5.

X. All else equal, higher X (and thus higher growth) depresses a firm's optimal size, as it is more likely to face creative destruction by its competitors.

In addition, notice that the intermediate good producers' optimal size is positively related to markups for $\phi > 1$. By taking ratios, we obtain a third relation between μ_L, μ_H, and h stated in the following Lemma[6]:

Lemma 3.4.

$$\frac{\alpha}{1-\alpha}\left(\frac{\frac{\mu_H^*-1}{\mu_H^*}}{\frac{\mu_L^*-1}{\mu_L^*}}\right)^{\frac{1}{\phi-1}} = \frac{h^*}{1-h^*} \tag{E3}$$

Along a balanced growth path, the three Equations E1, E2, and E3 define a system of three equations in three unknowns, h^*, μ_L^*, μ_H^* which are solved for numerically. Together, these values determine static productivity as shown below in Proposition 3.5.

To proceed, we take as given markups and the size distribution, and use those outcomes to solve for marginal costs for the respective firm types. We continue with the definition of the price index P together with the line level prices, $p_i = \gamma mc_{j'(i)}$, which depend on marginal costs:

$$P = \exp \int_0^1 \log \frac{p_i}{q_i} di = \frac{\gamma}{Q} \exp \int_0^1 \log mc_{j'(i)} di. \tag{39}$$

Plugging in the two firm types and using the price level normalization $P \equiv 1$, we get

$$Q/\gamma = (mc_H^*)^{h^*}(mc_L^*)^{(1-h^*)} = \left(\frac{\mu_L^*}{\mu_H^*}\right)^{h^*} mc_L^*, \tag{40}$$

which means we can solve for mc_L^*, mc_H^* in terms of Q^*. With marginal costs solved for, we recover equilibrium wages using Equation 35.

Using the Equation FOC$_j$ for both firm types together with the free entry condition yields a system of three equations in the three unknowns n_H^*, n_L^*, and X^*. The third equation in this system is derived from the firm-level profit shares:

$$\pi_j^* \equiv 1 - \frac{1}{\mu_j^*} \cdot \frac{\frac{\beta}{1-\sigma}}{1+\frac{\beta}{1-\sigma}} - \psi(X^*)^{\phi}(n_j^*)^{\phi-1}. \tag{41}$$

Recall that the firm-level revenue is $Y_t n_j^*$, we can therefore get rid of Y_t and express the

[6] See Appendix B for derivations.

somewhat simplified free entry condition as:

$$\zeta = \frac{\alpha \cdot n_H^* \cdot \pi_H^* + (1-\alpha) \cdot n_L^* \cdot \pi_L^*}{1-\rho}. \tag{FE}$$

For a given X, we solve for n_j, and verify whether $X = X^*$ using Equation FE. Notice that the labor supply elasticity facing firms $\frac{\beta}{1-\sigma}$ influences X^* directly through the free entry condition, and indirectly through its effect on equilibrium markups.

Since the mass of product lines is one, we can solve for the number of firms active in equilibrium:

$$J^* = \frac{1}{\alpha n_H^* + (1-\alpha) n_L^*}. \tag{42}$$

To solve for total output, we can make use of Equation 33 which contains the unknown equilibrium outcomes Y_t^* and z^*. Since z_t^* is a function of the only unknown Y_t^*, we search numerically for the Y_t^* that satisfies the equation. We then compute z_t^* using its definition which recovers L_j^* and therefore also aggregate employment, which is the last equilibrium outcome we need for characterizing aggregate output.

Aggregate output can now be characterized using a decomposition as in Boppart and Li (2021). A proof is given in Appendix B.4.

Proposition 3.5. *Aggregate output can be decomposed as follows:*

$$Y_t = Q_t \cdot S_t \cdot M_t \cdot L_t.$$

Where the first factor Q_t denotes the quality index and is given by:

$$Q_t \equiv \exp \int_0^1 \ln(q_{it}) di.$$

The second factor S^* measures aggregate process efficiency as a geometric average across product lines:

$$S_t \equiv \exp \int_0^1 \ln(s_{j(i,t)}) di.$$

The third factor M^* measures misallocation from employment dispersion which arise due to differences in line-level prices and process efficiencies:

$$M_t \equiv \frac{\exp \int_0^1 \ln \left(\frac{1}{mc_{j'(i,t)} s_{j(i,t)}} \right) di}{\int_0^1 \frac{1}{mc_{j'(i,t)} s_{j(i,t)}} di}.$$

The final factor L is aggregate employment and is defined as:

$$L_t \equiv \sum_{j \in J} L_j.$$

Starting with average process efficiency, S, we have that $S \equiv \exp \int_0^1 \ln(s_j(i))di$ which is a geometric average that is fully determined by the firm size distribution measured by h^*. It is therefore stationary along a balanced growth path and is given by $S^* = s_H^{h^*} s_L^{1-h^*}$.

With additional knowledge of firm-level marginal costs or markups, we can recover the misallocation measure M. This is the geometric mean of line level employment relative to the arithmetic mean of line level employment and therefore decreases as variance in employment across product lines increases. In our case, there are four combinations that arise from two possible leader- and follower-types. These outcomes depend on the share of the economy held by a given firm type, which is given by h^*. This outcome is also stable along a balanced growth path since balanced growth path marginal costs scale with Y_t^*. Intuitively, the factors S,M both boil down to the allocation of labor. For the former, this follows from labor allocation across firms, with production allocated toward unproductive firms. M on the other hand captures the effects arising from dispersion in prices, or the misallocation of labor across goods.

Since Q_t is growing along the balanced growth path, the contribution of quality on output can be described by the present value of future quality levels, $\{Q_t\}_{t=0}^{\infty}$, which is a geometric series given by $Q_0/(1-\rho g)$ that depends on an initial arbitrary quality level Q_0, and the growth rate g.

C. Competitive labor markets as a limit case

The model nests perfect competition in the labor market as a limit case when $\varepsilon = \beta/(1-\sigma) \to \infty$. Perfectly competitive labor markets can be recovered by either letting $\beta \to \infty$, in which case households' utility is infinitely sensitive to changes in consumption; or by letting $\sigma \to 1$, in which case households don't gain any utility from workplace amenities. In this section, we will discuss both cases and build intuition for the implications of the model. Detailed derivations are in Appendix A.7.

In the case when $\sigma \to 1$, the firms face a common labor supply curve:

$$\lim_{\sigma \to 1} L_j = \begin{cases} 0, & \text{if } \exists k \in \mathcal{J} : W_j < W_k \\ \mathcal{L}\left(\tilde{J}\left(1 + (\omega Y/W_j)^{\beta}\right)\right)^{-1}, & \text{if } W_j \geq W_k, \forall k \in \mathcal{J} \end{cases},$$

where $\tilde{J} \equiv \sum_{k \in \mathcal{J}} \mathbb{1}(W_k = W_j)$. Intuitively, if the wage of a given firm j, W_j is lower than at any other firm, all workers will go to that firm. Among all $\tilde{J}$ firms that pay the same wage, workers are split equally. That means there is a market wage $\bar{W}$, which is paid by any firm with positive size. The only choice margin that remains is between market labor at $\bar{W}$ and home labor at ωY.

When instead $\beta \to \infty$, the labor supply facing firm j is:

$$\lim_{\beta \to \infty} L_j = \begin{cases} 0, & \text{if } \exists k \in \mathcal{J} : W_j < W_k \vee \omega Y > W_j \\ \mathcal{L}(\tilde{J} + \tilde{J}^{\sigma})^{-1}, & \text{if } W_j \geq W_k \forall k \in \mathcal{J} \wedge \omega Y = W_j \\ \mathcal{L}(\tilde{J})^{-1}, & \text{if } W_j \geq W_k \forall k \in \mathcal{J} \wedge \omega Y < W_j. \end{cases}$$

Similarly, if $\beta \to \infty$ there is a market wage $\bar{W}$ which is paid by all firms that produce output. However, with $\beta \to \infty$, workers care about the wage much more than amenities or their relative preference over market labor and home production. A fraction or workers will chose the outside option if it pays the same as market labor, but not if it pays less.

Both cases imply that workers are indifferent between working for various firms. This means in equilibrium, a firm can hire as many workers L_j at market wage $\bar{W}$ as it wants, and the market wage $\bar{W}$ will be such that labor markets clear.

Turning to how this affects the model predictions, a common wage W implies that marginal costs are given by $mc_j = W/s_j$, which can also be derived from taking the limit of Equation 35. Limit pricing then implies line level markups are given by $\gamma s_{j(i)}/s_{j'(i)}$, and firm-level markups are given by $\mu_j = \gamma \cdot m \cdot s_j/W = \gamma \cdot s_j/(h^* s_H + (1-h^*)s_L)$.

In the present value decomposition, an increase in the share of H-type firms, h^*, under perfect competition increases the average efficiency S. At the same time, the misallocation from the employment dispersion, M, is increasing in h^* if $h^* < .5$, and decreasing otherwise. This is because having both types of firms decreases the average markup. Moreover, an increase in h^* implies that research is more costly due to convexity in the cost curve.

If $\gamma < \frac{s_H}{s_L}$, which is what we find in the data, $h = 1$, i.e. only the H-type firms produce output. This is not in line with the data, as firms exist which are less efficient in production. With perfect competition on the labor market, a firm distribution featuring active firms at different productivity levels hence requires $\gamma \geq \frac{s_H}{s_L}$. Introducing monopsony allows for a more flexible calibration, as the equivalent parameter restriction is

$$\gamma \geq \left(\frac{s_{j'}}{s_j}\right)\left(\frac{s_{j'}}{s_j}\frac{n_j}{n_{j'}}\right)^{\frac{1-\sigma}{\beta}}, \tag{43}$$

as derived in A.2. We will come back to this point in the quantitative part of this paper.

IV. Quantitative results

This section first describes the calibration strategy, including externally set parameters and targeted data moments. We also comment on a number of untargeted moments. To

Table 1: Targeted moments

Definition	Monopsony	PC	Data
Labor supply elasticity	7.13	1000	7.13
Unemployment rate	5.76%	5.76%	5.76%
Firms per capita J	0.11	0.11	0.11
Markup μ_H	1.29	1.29	1.29
Markup μ_L	1.14	1.14	1.14
Top 10% output share	75.92%	75.92%	75.92%
Growth rate	1.078%	1.078%	1.078%
Unemp. elasticity $\varepsilon_{u,\omega Y}$	0.10	0.10	0.02-0.32

"PC" refers to perfectly competitive labor markets, here approximated by setting the labor supply elasticity to 1000. Markup targets come from Edmond et al. (2023) Table 3, and correspond to their estimate of aggregate markup of 1.15, implied by the U.S.Census of Manufactures from 1972 to 2012. For μ_H we use the mean of reported markups for the 90th and 99th percentile of the markup distribution. For our main specification, we set $\varepsilon_{u,\omega Y} = 0.1$.

build intuition, we carry out the calibration for a case with competitive labor markets (denoted PC), and one with monopsony. Finally, we comment on how the model performs in matching untargeted moments and examine how outcomes change with the labor supply elasticity by varying σ.

A. Calibration

In this section, we describe the process used to calibrate the model's parameters. One set of parameters is assigned directly based on values commonly used in the literature or normalized. Another set of parameter values is obtained by using the equilibrium conditions of the previous section, combined with target values that allow us to obtain parameter values as residuals. This approach allows us to fit targeted moments exactly. We describe the procedure in the following text and summarize target moments and model outcomes in Table 1, and a summary is included in Appendix C.

For the first set of parameters, we assign them directly by normalizing or taking estimates from the literature. Normalized values include the initial quality level, $Q_0 = 1$; the process efficiency of the less process efficient firms, $s_L = 1$; and the population, $\mathcal{L} = 1$. We let $\rho = 0.95$, in line with commonly used values for the discount rate. We calibrate the labor supply elasticity facing firms, $\frac{\beta}{1-\sigma}$, to match the point estimate of a meta-regression of best-practice estimates of the elasticity reported by Sokolova and Sorensen (2021), Table 5. By centering our results around the emerging consensus estimate of the elasticity, our results speak to the aggregate effects of local variation in the elasticity. Finally, we define the share of highly process efficient firms as $\alpha = 0.1$. The choice of α affects choices for

Table 2: Calibrated parameter values

Parameter	Monopsony	Competitive	Description
ε	7.133	1000	Labor supply elasticity facing firms
α	0.100	0.100	Share H-type firms
ρ	0.950	0.950	Discount rate
s_L	1	1	Low process efficiency
s_H	1.682	1.135	High process efficiency
Φ	1.181	1.181	Research cost convexity
ψ	1.520	1.520	Research cost
$\mathcal{L}$	1	1	Population
γ	1.254	1.254	Innovation step
$\zeta \cdot (1-\rho)$	2.105	1.164	Entry cost
ω	$1.7 \cdot 10^{-12}$	$3.0 \cdot 10^{-12}$	Outside option
β	0.106	0.106	Utility sensitivity to consumption
$1-\sigma$	0.015	0	Utility sensitivity to workplace

calibration targets in what follows.

Applying our chosen value for α, we find the relative process efficiency of the two firm types, $\frac{s_H}{s_L} = 1.49$, in the data. We don't need to target this value and save it for model validation. From the same data source, we compute the average sales held by the α largest firms found in Compustat (Standard & Poor's, 2020) over the period 1954–2007 as $h = 75.92\%$.

Markups associated with the calibrated sales share h are obtained from Edmond et al. (2023), Table 3, and correspond to the column where their aggregate markup, estimated from the U.S. Census of Manufactures from 1972 to 2012, is 1.15. For μ_L, we use the median reported value. For μ_H, we use the mean of reported markups for the 90th and 99th percentiles of the markup distribution. The third equilibrium condition stated in Equation E3 now allows us to solve for ϕ.

Continuing to use the equilibrium conditions, we note that the first equilibrium condition stated in Equation E1 implies a value for the quality step parameter, $\gamma = h \cdot \mu_H + (1-h)\mu_L$, i.e., the output-weighted average markup. In turn, a value of γ means that a gross growth rate target, $g = 1.01078$, can be used to identify the equilibrium value of X using $\gamma^X = g$. In turn, X is used in the free entry condition, Equation FE, to solve for the entry cost parameter.

By additionally targeting the number of firms per worker found in Compustat (Standard & Poor's, 2020), we can use the first-order condition, Equation FOC$_j$, to solve for ψ.

To identify the level of σ and β, we use an estimate of the elasticity of unemployment

Table 3: Untargeted moments

Moment	Monopsony	PC
High-type Firms		
Labor share	0.680	0.774
Research share	0.090	0.090
Profit share	0.230	0.135
Low-type Firms		
Labor share	0.769	0.876
Research share	0.049	0.049
Profit share	0.181	0.074
Aggregate		
$\frac{W_H}{W_L}$	1.486	1.003
$\frac{s_H}{s_L}$	1.682	1.135

Wage and productivity numbers are taken from Compustat (Standard & Poor's, 2020).

with respect to the outside option, $\varepsilon_{u,\omega Y} = \beta(1-u)$, together with the unemployment rate. The elasticity of unemployment with respect to the outside option is challenging to estimate in the data. Nevertheless, there have been a number of natural experiments providing variation in unemployment insurance benefits that the literature has exploited to identify it. A good overview is found in Landais et al. (2018), who summarize estimates from such quasi-experimental settings. Most estimates are in the range of 0.02–0.32, and we take a middle stance, targeting 0.1 for our main specification. With β given, σ follows as a residual from the ratio $\frac{\beta}{1-\sigma}$.

The final parameter that needs to be calibrated is ω. This parameter follows from previous moments and equilibrium conditions, the idea is that we can obtain an expression for z without knowledge of ω. To obtain z, we start by calculating the relative firm-level labor using relative firm sizes, $s_H L_H/(s_L L_L) = n_H/n_L$, which gives firm-level labor from total employment:

$$L_L = \frac{L}{\alpha\frac{L_H}{L_L} + (1-\alpha)}, \quad L_H = \frac{L_H}{L_L} \cdot L_L. \tag{44}$$

With wages and firm-level labor, we can back out z using Equation 6 which describes the labor supply facing firms. We use the definition of z to back out ω. Y, which is another necessary equilibrium object inside z, is obtained from inverting Equation 32.

The calibrated model implies the untargeted moments summarized in Table 3. What stands out are the profit share, relative wages, and the relative process efficiencies. As mentioned previously, the competitive calibration fails to match the relative process

Table 4: Elasticities

	$d\log PVY$	$d\log S$	$d\log PVQ$	$d\log M$	$d\log L$
$d\log \frac{\beta}{1-\sigma}$	0.13	0.24	-0.14	0.03	0.00
$d\log \sigma$	8.76	16.01	-9.47	2.10	0.12
$d\log md$	1.08	1.97	-1.16	0.26	0.01

This table shows model elasticities around the calibrated value for β and σ. md refers to the markdown resulting from monopsony and is given by $md := \frac{\frac{\beta}{1-\sigma}}{\frac{\beta}{1-\sigma}+1}$.

efficiency measured in the data. This is because if $\gamma < \frac{s_H}{s_L}$, the L-type firms cannot produce when labor markets are competitive, as they pay the same wage and will end up without a markup sufficient to sustain profitable production. However, we also match a data moment with $h < 1$, as less process efficient firms have positive size in reality, in which case the model with competitive labor markets cannot match a realistic ratio of s_H/s_L. This demonstrates that monopsonistic labor markets are one channel that can rationalize the existence of firms with different constant process efficiencies, while also matching markup size and relative productivities.

Regarding the research and profit shares, the model seems to generate reasonable, although high, values when compared to aggregate measures. Looking at more granular data, such as damodaran2025margins for profit margins and Damodaran (2025) for R&D costs, suggests that these shares vary widely by industry. For example, the pre-tax profit share ranges from close to zero up to over 40%. Similarly, while there seems to be very little R&D spending in some industries, in others firms seem to invest almost half of their revenue in R&D. Our model outcomes match these shares roughly for industries such as electronics, healthcare products, and entertainment.

B. Comparative statics

Having calibrated a labor supply elasticity of $\varepsilon = \frac{\beta}{1-\sigma} = 7.133$, we vary σ holding other parameters constant. How a changing σ affect the components that make up aggregate output is shown in Figure 1. Key elasticities around the calibrated labor supply elasticity are shown in Table 4. We next turn to discussing the different forces driving these results. Towards the competitive extreme where $\sigma \to 1$, aggregate productivity increases and is the driving force of aggregate output. As the model moves away from imperfect competition on the labor market, economic activity reallocates towards more process efficient firms. As demonstrated in Figure 2, only more process efficient firms are active in perfectly competitive labor markets, as $\gamma < \frac{s_H}{s_L}$ according to the calibration. This condition implies that firms with a low process efficiency are not able to operate profitably. At the extreme distribution where $h^* = 1$, these firms do not produce at all. In turn, this implies uniform

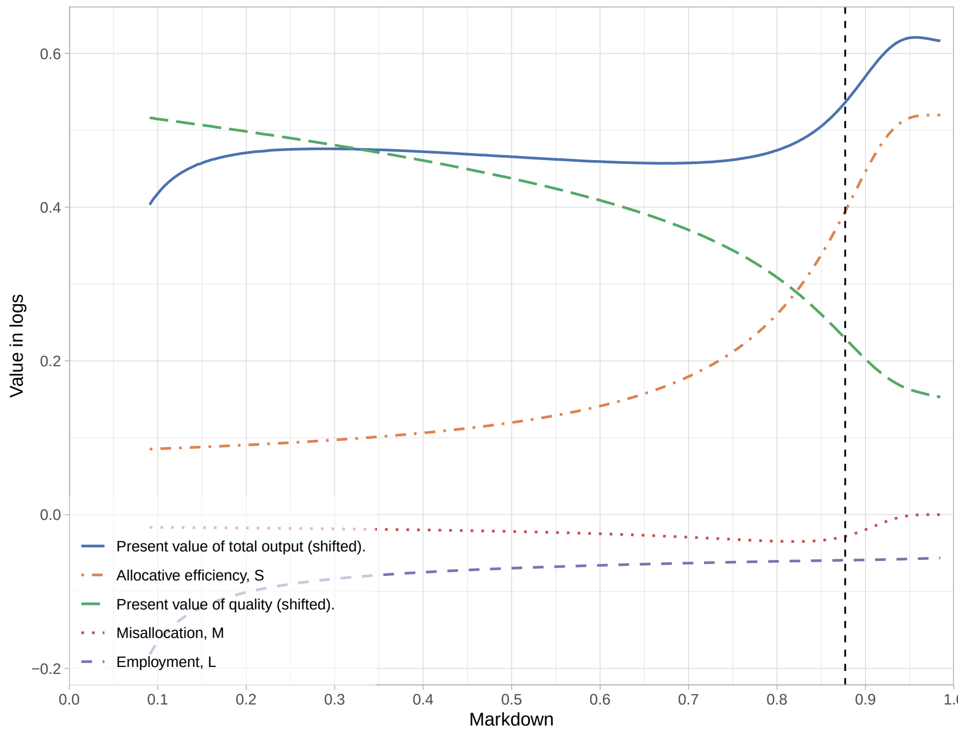

Figure 1: Decomposition comparative statics

Balanced growth path outcome for counterfactual values of σ. Effect on $\ln PV\{Y\} = \ln \frac{Q_0}{1-\rho g} + \ln S + \ln M + \ln L$ following changes in σ.

markups $\mu_H = \gamma$, i.e. no misallocation, $M = 1$, and that aggregate productivity $S = s_h$ is maximized, as shown in Figure 1.

In the case where $\gamma \frac{s_L}{s_H} > 1$, both firm types earn profits where they are quality leaders and are able to survive. The solution is therefore interior and we can use Equation E3 to find a value for h^*. What enables the less process efficient firms to survive under imperfectly competitive labor markets is their ability to keep relative marginal costs low by staying small enough so as to sustain a markup greater than 1. In other words, it is the ability of small firms to recruit employees with preferences for their specific work place that allows them to pay a lower wage and survive amidst competition of more process efficient firms.

That economic activity is reallocated towards the most process efficient firms is corroborated by Figure 2. The top left pane of the figure shows that the markup of intermediate producers with a low process efficiency decreases towards one. Meanwhile, the same figure shows that markups for the most process efficient firms initially increase with σ, until they eventually decrease toward γ. The non-monotonic relation is the result of stable output prices, and two forces that pull the marginal costs in opposite directions. First, increasing the labor supply elasticity facing firms is equivalent to decreasing the elasticity of wages with respect to size, meaning that wages increase slower with size and thereby

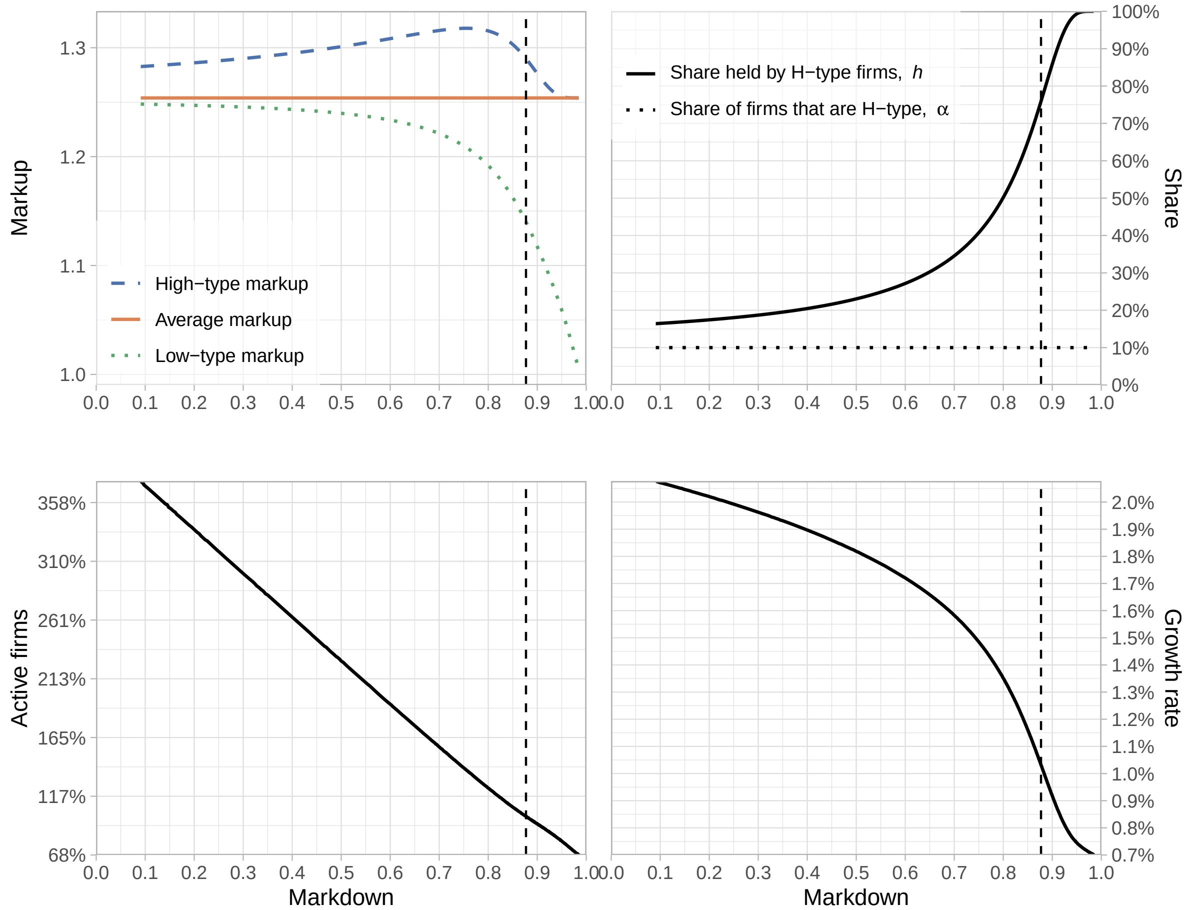

Figure 2: Outcomes for varying σ.

Note: The markdown is given by $\frac{\frac{\beta}{1-\sigma}}{1+\frac{\beta}{1-\sigma}}$. The vertical line indicates the point to which the model was calibrated.

create less upward pressure on marginal costs. Second, reallocation of economic activity towards the most process efficient firms raises wages, thereby increasing marginal costs. Since the harmonic mean, m, of marginal costs tends to be stable, the expected output price given by γm is stable as well. This results in declining markups for both firm types with $\sigma \to 1$. This within-firm effect is mostly compensated by the between-firm effect of reallocation toward high markup firms, (sales-weighted) average markups thus barely move, implying that the profit share stays largely constant.

Quantitatively, the second strongest force is the effect on growth. As $\sigma \to 1$, a producer's marginal product line becomes less profitable due to declining markups. For low values of σ, this effect is additionally related to declining entry, see Figure 2, bottom left pane. Over the range of σ we consider, the two main effects work in the same direction resulting in a monotonic decline of the contribution of growth on the present value of total output that is quantitatively weaker than the effect of reallocation for $\sigma \to 1$.

The third most important contribution is driven by increasing employment. While the number of firms are declining, leaving households with fewer workplace options, the increasing wages fully offset this effect across the distribution of σ.

Finally, misallocation from dispersion in the marginal revenue product plays a minor role in determining aggregate output.

The combined effect is that aggregate output is increasing nearby our calibrated value of σ, and exhibits a non-monotonic shape due to competing forces that vary in their respective strength.

C. Discussion

Labor market power affects economic growth via two main channels in this model. The first is the effect on the profit share, as greater labor market power leads to a larger profit share and an increase in the number of active firms. Since innovation is cheaper when done by small firms, the overall amount of research and growth increases with labor market power. The second effect is the effect that labor market power has on the market structure and markups. Around the calibrated value for labor market power, more monopsony increases the markup for each producer type. This is because it allows less productive firms with high marginal costs to expand, making it more likely that a given follower has a high marginal cost. With limit pricing, this increases intermediate product prices. While both firm types observe greater markups, the economy-wide average markup stays roughly constant due to reallocation towards low markup firms.

To separately identify the quantitative importance of the two aforementioned effects, Figure 3 shows two counterfactuals. First, firm entry is shut down which isolates the effect of changing markups and market structure on growth. Second, markups are held constant

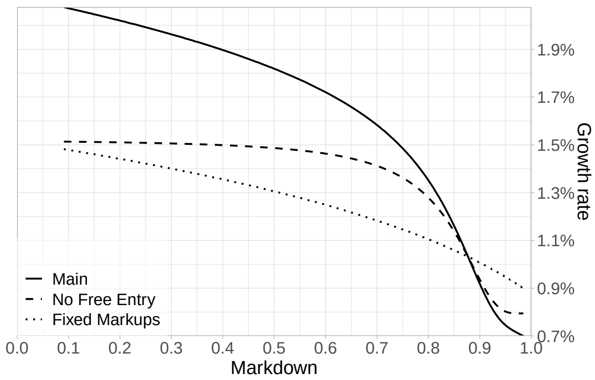

Figure 3: Decomposition of growth mechanisms.

The graph shows the effect of shutting down free entry and markup variation on growth. Around the calibrated value where curves intersect, disabling free entry preserves most variation in growth, implying that markup variation drives the results.

at their calibrated value such that firm entry is the driver of growth. Jointly, these two effects explain virtually all variation in growth. [7] Quantitatively, the effect that labor market power has on market structure and markups accounts for nearly all variation in growth near the calibrated value.

These results emphasize that around the calibrated equilibrium, the main driver of growth as a response to shifts in labor market power are changing markups. This implies that the interaction of product market power and labor market power is a key component in understanding growth effects. While product market power applies within a given product line, labor market power spans various product lines and therefore affects innovation incentives of the firm through its effect on the marginal cost distribution.

In practice, these results apply to economies where firms face a finite labor supply elasticity, warranting a comment on potential determinants of this key elasticity. One mechanism that leads to monopsony is migration behavior. Where migration is costly, workers living further away will require being paid a compensating wage. At the same time, policies such as working from home make it easier for firms to hire a large number of workers without paying much higher wages. An even more direct effect of policy works through minimum wages. These imply that up to a certain firm size, the labor supply elasticity is infinite.

V. Extension: Progressive income taxation

In this section, we extend the model to incorporate progressive income taxation. The previous section demonstrated that introducing a finite firm-level labor supply elasticity affects economic aggregates. Here, we examine how tax policy can alter this elasticity.

[7] There is a third effect driven by how changing markups affect profit shares. This effect is small since the expected markup of an entrant stays roughly constant as the reallocation effect cancels the within-firm rise in markup.

Specifically, we show that progressive income taxes influence the labor supply elasticity with respect to a firm's gross wage when the elasticity is finite, in line with Berger et al. (2024). After describing the extension, we turn to a discussion of how the extended model can be applied to evaluate tax reforms, such as those under the Reagan administration. Refer to Appendix D for technical details.

We begin by assuming that a government raises revenue for government consumption from a wage bill tax at the firm-level[8]. The government provides each household with G_t units of a public good. Household preferences are then given by:

$$u_{ojt} = \beta \ln C_{ojt} + \xi_{ogt} + (1-\sigma)\epsilon_{ojt} \tag{45}$$

$$= \underbrace{\beta\eta}_{\tilde{\beta}} \ln(W_{jt}) + \beta(1-\eta)\ln(G_t) + \xi_{ogt} + (1-\sigma)\epsilon_{ojt}, \tag{46}$$

where total consumption C_{ojt} is a Cobb-Douglas aggregate of private (C^p_{ojt}) consumption paid for with wage income, and government consumption (G_t) enjoyed by the household: $C_{ojt} = W^{\eta}_{jt} G^{1-\eta}_t$. The modified utility results in labor supply similar to the formulation in the base model, with the labor supply curves increasing in the net wage offered by a firm:

$$L_j(W_{jt}) = z_t W_{jt}^{\frac{\tilde{\beta}}{1-\sigma}}. \tag{47}$$

Market clearing now features a term for the tax-financed government consumption, given by $\mathcal{L}G_t$, and reads as follows:

$$Y_t = \tilde{C}_t + C^p_t + \mathcal{L}G_t + E_t, \tag{48}$$

where private consumption is the sum of all net wages paid, and government expenditure is the sum of all taxes paid:

$$C^p_t = \sum_{jt} W_{jt} L_{jt}, \quad \mathcal{L}G_t = \sum_j T(W_{jt}/\bar{W}_t) W_{jt} L_{jt}. \tag{49}$$

We assume that the government commits to a tax schedule $T(W_j/\bar{W})$, where $\bar{W}_t \equiv \sum_{j=1}^{\mathcal{J}} W_{jt}L_{jt} / \sum_{j=1}^{\mathcal{J}} L_{jt}$ is a reference wage. The government's objective is then simply to spend such that its budget constraint binds in each period.

With the income tax levied in full on the firm, we obtain a formulation that shows how progressive taxes directly impact a firm's marginal cost. Due to the crucial role of marginal costs in the model, this will feed through into markups, research decisions and aggregate

[8] This is equivalent to raising taxes on the worker side, for details refer to Appendix D.1

outcomes. The full firm problem becomes:

$$\Pi_j(Y_t, \{q_{ijt}\}_{i\in[0,1]}) = \max_{\{p_{ijt}\}_{i\in[0,1]}} \int_0^1 p_{ijt} y_{it} di - (1 + T(W_{jt}/\bar{W}_t)) W_{jt} L_{jt}, \tag{50}$$

$$\text{s.t.} \quad L_{jt} = \int_0^1 f^{-1}(y_{ijt}) di, \quad W_{jt} = \left(\frac{L_{jt}}{z_t}\right)^{\frac{1-\sigma}{\tilde{\beta}}} \tag{51}$$

$$\text{Intermediate product demand as in (3).} \tag{52}$$

We now define the tax rate following Borella et al. (2022). This is in line with standard formulations in the literature, as used by Berger et al. (2024). For simplicity, we reformulate this set-up such that (1) the tax transaction is paid by the firm rather than the worker[9] and (2) the tax rate is based on the wage relative to the average net wage, not median gross wage. The average net wage $\bar{W}_t$ can be interpreted as a reference wage, or the average wage employed workers receive. The parameter λ, similar to Borella et al. (2022), governs the base level of the tax. More importantly for us, τ pins down income tax progressivity, with $1-\tau$ being the elasticity of post tax income w.r.t. pretax income. The tax schedule is represented by:

$$T\left(\frac{W_{jt}}{\bar{W}_t}\right) = \left(\frac{W_{jt}^{\tau}}{1-\lambda}\frac{1}{\bar{W}_t^{\tau}}\right)^{\frac{1}{1-\tau}} - 1, \quad \bar{W} = \frac{J_h L_{ht} W_{ht} + J_l L_{lt} W_{lt}}{J_h L_{ht} + J_l L_{lt}}. \tag{53}$$

A key channel through which income tax progressivity acts in the model is then through the firm-level labor supply elasticity with respect to the gross wage. In addition to an upward-sloping labor supply curve due to idiosyncratic preferences, the elasticity now has a component resulting from tax policy:

$$\frac{\partial \log(L_j)}{\partial \log(1 + T(W_{jt}/\bar{W}_t)) W_{jt}} = \underbrace{\frac{\tilde{\beta}}{1-\sigma}}_{\text{Preferences}} \underbrace{(1-\tau)}_{\text{Policy}}. \tag{54}$$

Through this elasticity, tax progressivity influences aggregate variables similarly to the discussion in Section B. In the next section, we apply the model to a historic tax reform and quantitatively solve the model with taxes in order to discuss the effect of progressivity.

A. Policy application: 1980s tax reform

We now turn to a policy application of the extended model. The goal is to gain a quantitative understanding of the importance of the firm-level labor supply elasticity when considering changes to tax policy. As shown in the previous section, the firm-level labor

[9] Note that this tax set-up is equivalent to one where the tax is levied on the worker as shown in Appendix D.1.

supply elasticity is affected by tax policy. The way in which taxes are raised therefore matters for the degree of monopsony: changing the average tax level λ has no effect on the firm-level labor supply elasticity, whereas changing the income tax progressivity through τ does.

Building on the estimation results by Borella et al. (2022), the 1980s tax cuts under the Reagan administration reduced both the average level and the progressivity of income taxes. Table 5 summarizes key changes to tax policy. Equation 54 then suggests the firm-level labor supply elasticity would have increased by about 2.3%.

Parameter	1981	1988
Tax average λ	0.113	0.096
Tax curvature τ	0.084	0.063

Table 5: Parameter changes in tax reform

In order to compare the state of the economy before and after the reform, we consider the tax cuts in isolation as if they were a single reform which brought the economy from one balanced growth path (pre-1981) to another (post-1988). Since the underlying tax rates on real incomes fluctuate significantly due to inflation and irregular reforms, the approach serves not to evaluate the reform per se, but rather to gain a quantitative notion of how monopsonistic labor markets might have contributed to the final outcome. For this purpose, Table 6 lists model outcomes next to observed outcomes. Note that the increase in output from the model is due to both increases in static efficiency, and the (largely unaffected) rate of productivity growth. The "no reform" row refers to the model prediction of the changes between 1981 and 1987 without the tax reform, i.e. remaining on the previous balanced growth path.

Definition	TFP growth	Growth in Q	Growth in S
Data	7.74%	N/A	N/A
Model with reform	7.45%	6.85%	0.56%
Model without reform	6.93%	6.93%	0.00%

Table 6: Tax reform: Changes 81 – 87

Note: Data, model and no reform values are all changes in percentages.

Our comparison between model and data is based on an annual productivity time series from BLS (2024) data. Despite simply adding the taxation on top of the calibration described in the previous section, the model matches overall TFP growth in the period of study quite well. We find that the reform did lead to an increase in the post-period productivity. However, this came at the cost of a slightly decreased rate of quality growth.

Macnamara et al. (2024) argue that income tax cuts may boost growth in the long run, but our model suggests that a lower rate of income tax progressivity may reduce growth by increasing competition and stifling incentives to innovate.

B. *Alternative tax schemes*

In this exercise, we consider a budget-neutral tax reform, i.e. we fix the detrended level of government spending G at the BGP level. We then consider those combinations of average tax rate, λ, and progressivity, τ, that generate the same revenue for the government, and look at outcomes along that path. Around its calibrated value, a 1% decrease in τ increases the labor elasticity by approximately 0.1%.

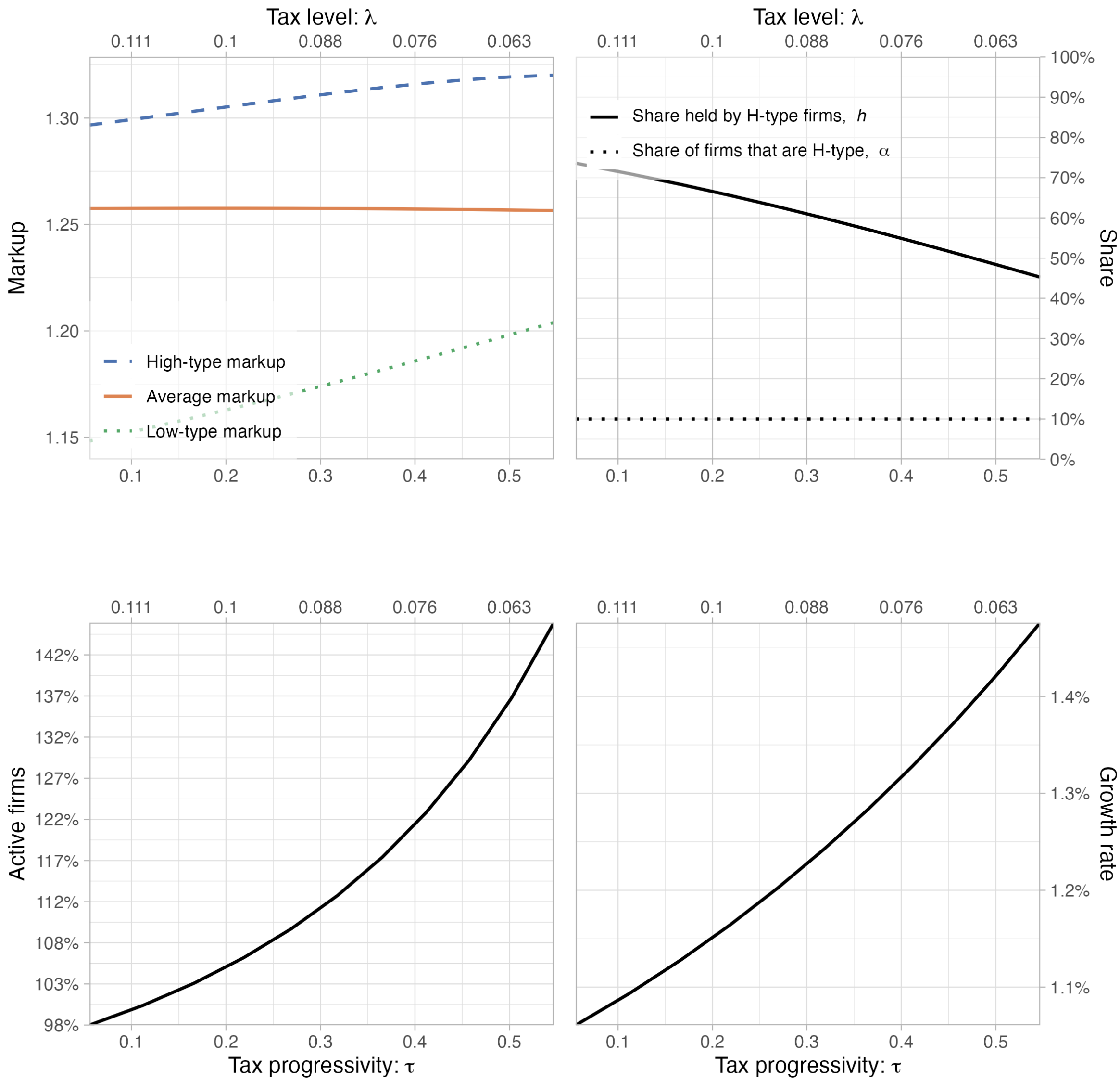

Figure 4: Outcomes at different (λ, τ), fixed G

Results from a budget-neutral tax reform. In each graph, tax progressivity τ decreases from left to right, while the tax level λ increases. The number of firms is relative to base calibration.

The effects here are mainly driven by the change in τ, which can be seen in Figure 4. The increased labor supply elasticity in from lowering τ increases concentration as process efficient H-type firms expand. As in the previous section, this will increase static productivity, as more production labor is allocated to firms with high process efficiency. However, it will also affect the long-run growth rate negatively as research is less efficient at larger firms, and lower markups additionally make innovation less valuable to the firm.

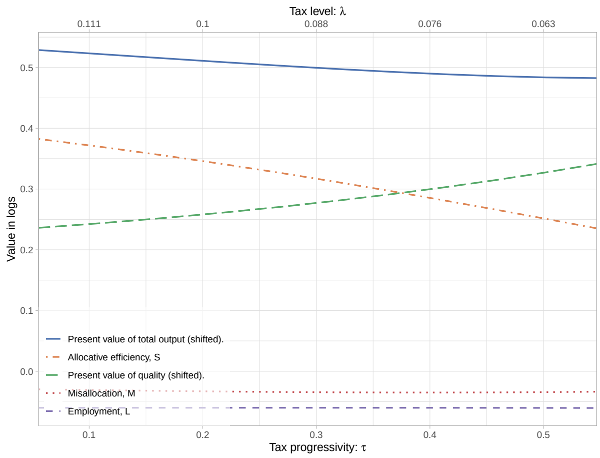

Figure 5: PV decomposition

All components are relative to their values in the base calibration, values at different (λ, τ), fixed G.

As in Section 4, the effect on the present value of output can be decomposed into the different channels from Proposition 3.5. Figure 5 shows the present value of output per capita and its components. Similarly to the change in σ considered in the previous section, the main trade-off is between static productivity and long-run growth. Again, the high elasticity (low τ) case maximizes the net present value of output at the cost of long-run growth.

VI. Conclusion

This paper explores the role of monopsony power in a growth model with product market power and creative destruction. For one, we find that this allows matching data moments models of this class may otherwise be unable to fit. Primarily, our main finding is that the presence of monopsonistic labor markets implies a trade-off between current output and long-run growth. Monopsony power reallocates labor towards smaller, less productive firms, which reduces static output, while simultaneously increasing invectives to innovate. The paper further shows that the interaction between labor market power and product market power is important for understanding firms innovation incentives. In addition, labor market power increases the profit share, which stimulates entry which can drive growth if innovation is done more efficiently at a small scale, as labor market power directs resources towards smaller firms. We also demonstrate that the monopsony-induced finite labor supply elasticity is policy-relevant. We show that, to some extent, it is directly affected by the level of tax progressivity. We quantify the relationship between monopsony power, income taxation, and aggregate outcomes. Our model illustrates how monopsony power, and tax progressivity, influence labor allocation, output, and research efficiency. Broadly, our findings suggest that monopsony power should be considered in the context of economic growth. To find an 'optimal' level of monopsony which maximizes discounted output, the trade-off between static and dynamic efficiency is crucial. In The framework we develop also serves as a foundation for future investigations to determine this optimal level. Minimum wages could be an avenue for future research. In our set-up, minimum wages would force low productivity firms to decide between operating at a higher wage – or not at all. A minimum wage thus potentially drives these firms out of the market and affects the shape of the labor supply curve. Similarly to changes in labor supply elasticities, this would then affect both aggregate productivity and growth.

APPENDIX A. Model derivations

A.1 Intermediate good demand from CD aggregator

The optimization problem of the final goods producer is:

$$\max_{\{y_i\}_{i\in[0,1]}} P\left(\exp\int_0^1 \ln(q_i y_i)di\right) - \int_0^1 p_i y_i di \tag{55}$$

The first order condition is given as: $P\exp\int_0^1 \ln(q_i y_i)di\frac{1}{q_i y_i}q_i - p_i = 0 \Leftrightarrow PY = p_i y_i$ The price index P is given as: $Y = \exp\int_0^1 \ln\left(q_i PY/p_i\right)di \Leftrightarrow P = \exp\int_0^1 \ln(p_i/q_i)di$.

A.2 Nash equilibrium in the pricing game

The equilibrium concept we use to solve the model is Bertrand, which means that within each differentiated intermediate good market, firms set prices to maximize profit taking all other firms' prices as given. Within a product line, goods are assumed to be perfect substitutes, which means the firm that posts the lowest quality-adjusted price attracts all demand $p_i y_i = PY$. Note that posted prices are binding and whoever attracts demand produces to fulfill that demand.

Note that we assumed that the final goods producer only buys one (quality) type of each variety. This requires two assumptions: (i) if the quality-adjusted prices are equal, the final goods producer prefers the higher quality product, and (ii) prices posted by the quality leaders are always greater than their respective marginal cost.

The first assumption allows us to rule out collusive equilibria wherein firms split product markets at a collusive price.

The second condition holds for a sufficiently large quality step size, $\gamma > \frac{mc_k}{mc_l}, \forall k, l \in \{1, ..., \mathcal{J}\}$, or:

$$\gamma \geq \left(\frac{s_{j'}}{s_j}\right)\left(\frac{s_{j'}}{s_j}\frac{n_j}{n_{j'}}\right)^{\frac{1-\sigma}{\beta}} \tag{56}$$

In one equilibrium of the model, the quality leader sets the quality-adjusted price equal to his follower's quality-adjusted marginal cost and the quality follower sets its price equal to its marginal cost. [10] Under this pricing, the quality adjusted prices of the firms are equal. Due to the tie-breaking rule, the quality leader produces the full demand for products in the given product line i and the follower produces nothing and earns zero surplus.

The price in a given product line is thus given by $\frac{p_{j(i)}}{q_{j(i)}} = \frac{mc_{j'(i)}}{q_{j'(i)}} \Leftrightarrow p_{j(i)} = \gamma mc_{j'(i)}$, where $j'(i)$ indexes the 'follower' in a given market i, and $j(i)$ the quality leader.

[10] There is a continuum of Nash equilibria where the quality leader posts a lower price and followers posts prices below their marginal cost.

The follower has no profitable deviation, since lower prices imply selling below marginal cost, and higher prices generate no sales. Meanwhile, there is no profitable unilateral deviation by the quality leader since a higher price loses all demand, and a lower price reduces the price without affecting output.

A.3 Labor supply: Nested discrete choice

The labor supply choice is modeled as a nested discrete choice problem. For technical details of the derivation of choice probabilities, refer to Train, 2009 or McFadden, 1977. In the following we drop time subscripts for wages for readability. The set up is described in B.

It is possible to solve the household problem in two steps. First, conditional on choosing to work, the household chooses their preferred employer j^*. This can be written as: $j^* : \beta \ln(W_{j^*}) + \xi_{oe} + (1-\sigma)\varepsilon_{oj^*} \geq \beta \ln(W_k) + \xi_{oe} + (1-\sigma)\varepsilon_{ok}, \quad \forall k \in \mathcal{J}$. Since ξ_{oe} is shared across jobs, these terms drop out. Further, we can divide by $(1-\sigma)$ and get: $j^* : \varepsilon_{oj^*} - \varepsilon_{ok} \geq \frac{\beta}{1-\sigma} \ln(W_k/W_{j^*})$

The conditional choice probability for firm j^* is then given as:

$$p_{j^*,e} = P\left(\varepsilon_{oj^*} \geq \varepsilon_{ok} + \frac{\beta}{1-\sigma} \ln(W_k/W_{j^*}), \forall k \in \mathcal{J}\right) = \frac{W_{j*}^{\frac{\beta}{1-\sigma}}}{\sum_{k\in\mathcal{J}} W_k^{\frac{\beta}{1-\sigma}}}. \tag{57}$$

The last equality follows from the assumption on the distribution of ε_{oj}.

Further, the probability of choosing any job over home production is given as:

$$p_{g=e} = \frac{\left(\sum_{k\in\mathcal{J}} W_k^{\frac{\beta}{1-\sigma}}\right)^{1-\sigma}}{\left(\sum_{k\in\mathcal{J}} W_k^{\frac{\beta}{1-\sigma}}\right)^{1-\sigma} + ((\omega Y)^{\frac{\beta}{1-\sigma}})^{1-\sigma}} \tag{58}$$

The unconditional choice probability of choosing a given employer is equal to the product of the probability of choosing employment, $g = e$, and the conditional probability of choosing firm j given $g = e$:

$$p_{j*} = p_{g=e} * p_{j^*,e} = \frac{W_{j*}^{\frac{\beta}{1-\sigma}}}{\left(\sum_{k\in\mathcal{J}} W_k^{\frac{\beta}{1-\sigma}}\right) + (\omega Y)^{\beta} \left(\sum_{k\in\mathcal{J}} W_k^{\frac{\beta}{1-\sigma}}\right)^{\sigma}} \tag{59}$$

We define:$D_e \equiv \sum_{k=1}^{\mathcal{J}} W_k^{\frac{\beta}{1-\sigma}}$, $z \equiv \frac{\mathcal{L}}{D_e^{\sigma}(\omega Y)^{\beta} + D_e}$. To get to the labor supply $L_j(W_j)$ facing firm j, we simply multiply by the mass of households in the economy, $\mathcal{L}$, and use the definitions to get:$L_j(W_j) = zW_j^{\frac{\beta}{1-\sigma}}$. The labor supply elasticity faces by the firm is:$\frac{\partial L_j/L_j}{\partial W_j/W_j} = \frac{\beta}{1-\sigma}$.

A.4 Elasticity of the outside option

Note that the mass of workers that choose the outside option is given as: $U_t = \mathcal{L}\frac{(\omega Y)^\beta}{(\omega Y)^\beta + (\sum_{k=1}^{\mathcal{J}} W_{kt}^{\frac{\beta}{1-\sigma}})^{1-\sigma}}$. From here, the unemployment rate is found by dividing by $\mathcal{L}$. Taking the derivative w.r.t. (ωY) then gives: $\varepsilon_{u,\omega Y} = \frac{\partial u}{\partial(\omega Y)}\frac{(\omega Y)}{u} = \beta(1-u)$.

A.5 Capital demand

Firms make profits that are paid out to firm owners via interest rates, and therefore discount future profits at rate $\frac{1}{1+r_t}$. Firms are owned by capitalists, who allocate consumption and investment to maximize lifetime utility.

$$\max_{\{c_t,k_{t+1}\}_{t=0}^{\infty}} \rho^t \ln(c_t) \quad s.t. k_{t+1} = (1+r_t)k_t - c_t \tag{60}$$

Taking first order conditions yields the Euler equation $\frac{c_{t+1}}{c_t} = (1+r_t)\rho$, and since consumption grows at a constant rate on a BGP, we have a constant interest rate $r^* = \frac{g_c}{\rho} - 1$, where g_c is the growth rate of consumption, and $R^* = \frac{\rho}{g_c}$.

A.6 Marginal cost

Marginal costs of production, as relevant for the intermediate product market competition, is derived from the following cost function: $C(Y_{jt}) = W(L(Y_{jt}))L(Y_{jt})$, $L(Y_{jt}) = \frac{Y_{jt}}{s_j}$, $W(L) = \left(\frac{L}{z}\right)^{\frac{1-\sigma}{\beta}}$.

We get $mc_{jt} \equiv C'(Y_{jt}) = \frac{1+\frac{\beta}{1-\sigma}}{\frac{\beta}{1-\sigma}}\left(\frac{Y_{jt}}{s_j z_t}\right)^{\frac{1-\sigma}{\beta}}\frac{1}{s_j}$. Moreover, $Y_{jt} = \frac{n_{jt}Y_t}{\gamma m_t}$, and therefore: $\frac{\partial C_{jt}}{\partial n_{jt}} = mc_{jt}\frac{Y_t}{\gamma m_t}$.

A.7 Perfectly competitive labor markets

In the case when $\sigma \to 1$, the size of a firm paying wage W_j, is given by the labor supply equation given by Equation 6, evaluated in the limit. First, rewrite the labor supply equation by dividing through the numerator. Then, we distinguish different cases:

$$\lim_{\sigma\to1} L_j = \lim_{\sigma\to1} \mathcal{L}\frac{1}{\left(\sum_{k=1}^{\mathcal{J}}\left(\frac{W_{kt}}{W_{jt}}\right)^{\frac{\beta}{1-\sigma}}\right)^{\sigma}\left(\frac{\omega Y}{W_{jt}}\right)^{\beta} + \sum_{k=1}^{\mathcal{J}}\left(\frac{W_{kt}}{W_{jt}}\right)^{\frac{\beta}{1-\sigma}}}$$

We can distinguish three cases to evaluate the expression $(W_{kt}/W_{jt})^{\frac{\beta}{1-\sigma}}$ in the limit for $\sigma \to 1$:

$$\lim_{\sigma\to1}(W_{kt}/W_{jt})^{\frac{\beta}{1-\sigma}} = \begin{cases} 0, & \text{if } W_k < W_j \\ 1, & \text{if } W_k = W_j \\ \infty, & \text{if } W_k > W_j \end{cases}$$

Having made this observation, the labor supply in the limit is as in section C.

A.8 Growth rates

Consider a balanced growth path for which all growth stems from quality improvements, i.e. $\frac{Q'}{Q} = \frac{Y'}{Y} = g$. Assume that n_j is constant on BGP, and so is the number of firms and the revenue share of H-type firms, h. First, consider wage growth: $g_w = \frac{(\frac{Y'}{m'z'})^{\frac{1-\sigma}{\beta}}}{(\frac{Y}{mz})^{\frac{1-\sigma}{\beta}}} = \left(\frac{g}{g_m g_z}\right)^{\frac{1-\sigma}{\beta}}$

The growth rate of the marginal cost index is defined as: $g_m \equiv \frac{(\frac{h}{mc'_H} + \frac{1-h}{mc'_L})^{-1}}{(\frac{h}{mc_H} + \frac{1-h}{mc_L})^{-1}}$.

As this growth rate has to be constant on a BGP, it must hold that marginal costs of either firm type grow at the same rate, and the marginal cost index m grows at that rate as well: $g_m = \frac{mc'_L}{mc_L} = \frac{mc'_H}{mc_H}$

So we can look to either firm type j to figure out the growth rates: $g_m = \frac{mc'_j}{mc_j} = \frac{W'_j}{W_j} = g_w$. Next, the growth rate of z is defined as follows:

$$g_z = \frac{z'}{z} = \frac{[(g_w^{\frac{\beta}{1-\sigma}} \sum_k W_k^{\frac{\beta}{1-\sigma}})^\sigma g^\beta (\omega Y)^\beta + g_w^{\frac{\beta}{1-\sigma}} \sum_k W_k^{\frac{\beta}{1-\sigma}}]^{-1}}{[(\sum_k W_k^{\frac{\beta}{1-\sigma}})^\sigma (\omega Y)^\beta + \sum_k W_k^{\frac{\beta}{1-\sigma}}]^{-1}}$$

From this we can see that, constant g_z requires: $g_w^{\frac{\sigma\beta}{1-\sigma}} g^\beta = g_w^{\frac{\beta}{1-\sigma}} \Leftrightarrow g_w = g$. Combining the growth rates from wages and z we get: $g_z = g^{-\frac{\beta}{1-\sigma}}$ and $g = g_Y = g_Q = g_m = g_w = \gamma^X$.

APPENDIX B. Proofs of Lemmata and decomposition

B.1 Proof Lemma 3.1

Proof. Multiply both sides of $m_t^{-1} \equiv \frac{h_t}{mc_{Ht}} + \frac{1-h_t}{mc_{Lt}}$ with γm_t, note the definition of markup $\mu_{jt} = \frac{\gamma m_t}{mc_{jt}}$ and rearrange:

$$\gamma m_t m_t^{-1} = \frac{h_t \gamma m_t}{mc_{Ht}} + \frac{(1-h_t)\gamma m_t}{mc_{Lt}} \Leftrightarrow \gamma = h_t \mu_{Ht} + (1-h_t)\mu_{Lt} \Leftrightarrow -\frac{\gamma - \mu_{Lt}}{\gamma - \mu_{Ht}} = \frac{h_t}{1-h_t}$$

□

B.2 Proof Lemma 3.2

Using the expression for the optimal wage, which is an outcome of the static firm choice, $W_{jt} = s_j \cdot mc_{jt} \cdot \frac{\frac{\beta}{1-\sigma}}{1+\frac{\beta}{1-\sigma}}$, together with the identities $n_H \alpha J = h$ and $n_L(1-\alpha)J = (1-h)$, the expression in Lemma 3.1 is derived as follows.

Proof.

$$\frac{s_H}{s_L} \frac{mc_{Ht}}{mc_{Lt}} = \left(\frac{\frac{n_{Ht}}{s_H}}{\frac{n_{Lt}}{s_L}}\right)^{\frac{1-\sigma}{\beta}} \implies \left(\frac{s_H}{s_L}\right)^{\frac{\beta}{1-\sigma}+1} \cdot \left(\frac{\mu_L}{\mu_H}\right)^{\frac{\beta}{1-\sigma}} = \frac{h(1-\alpha) \cdot J}{\alpha(1-h) \cdot J}$$

This can be rearranged to conclude the proof. □

B.3 Proof Lemma 3.4

Proof. From $\frac{n_H}{n_L}$, using Equation FOC$_j$ cancels out the aggregate variables, which yields $\frac{n_H}{n_L} = \left(\frac{\frac{\mu_H - 1}{\mu_H}}{\frac{\mu_L - 1}{\mu_L}}\right)^{\frac{1}{\phi-1}}$. Moreover, by definition of h_t, the following holds: $\frac{h\cdot(1-\alpha)}{(1-h)\cdot\alpha} = \frac{n_H}{n_L}$. Setting the two expressions for n_H/n_L equal concludes the proof. $\square$

B.4 Proof for 3.5

Proof.

$$Y \equiv \exp\int_0^1 \ln(q_i y_i)di = \exp\int_0^1 \ln(q_i)di \cdot \exp\int_0^1 \ln(s_{j(i)})di \cdot \exp\int_0^1 \ln l_i di$$

$$= Q\cdot S\cdot \frac{\exp\int_0^1 \ln l_i di}{\int_0^1 l_i di}\cdot \int_0^1 l_i di = Q\cdot S\cdot \frac{\exp\int_0^1 \ln\left(\frac{Y}{\gamma mc_{j'(i)}s_{j(i)}}\right)di}{\int_0^1 \frac{Y}{\gamma mc_{j'(i)}s_{j(i)}}di}\cdot \sum_{j\in J} L_j = Q\cdot S\cdot M\cdot L$$

$\square$

APPENDIX C. Additional details quantitative part

Calibration summary:

- **Externally set parameters:** α, ρ
- **Normalizations:** $s_L = 1$, $Q_0 = 1$, $P = 1$, $\mathcal{L} = 1$
- **Calibrated parameters:** $s_H, \gamma, \psi, \phi, \omega, \sigma, \beta, \zeta$, and $\epsilon = \frac{\beta}{1-\sigma}$
- **Moments:** u, $\frac{L}{\mathcal{L}}$, $\frac{s_H}{s_L}$, $\mu_H, \mu_L, h, g, \varepsilon_{u,\omega Y}$

Compustat data: We use compustat data (Standard & Poor's, 2020) from 1954 to 2016, and take averages of various time periods for different applications. We focus on firms in the U.S. manufacturing sector by filtering the dataset to include only firms under NAICS codes starting with '31', '32', or '33', and reporting in U.S. dollars. Missing values were addressed by excluding firms without key variables like sales and employment, and only firms with positive sales and employment values were kept. Firms were categorized annually into the top 10% by sales and the remaining 90%. We calculate two key metrics: the revenue share of the top 10%, and the relative average revenue per employee, which compares the production efficiency of the top 10% with the bottom 90%.

APPENDIX D. Extension with wage taxes

D.1 Equivalence of tax setups

This is a brief note on the equivalence of two tax set-ups: (i) the firm pays a wage bill tax T^f on the net wage the worker receives, W^w, or (ii) the worker pays a wage tax T^w on the gross wage the firm pays, W^f. The relationship between the gross and net wages under the two tax regimes are summarized as follows: $W^f = W^w(1 + T^f(W^w))$

and $W^w = W^f(1 - T^w(W^f))$ Rearranging then yields a mapping between the two tax schedules: $T^w(W^f) = 1 - 1/(1 + T^f(W^f(1 - T^w(W^f))))$.

D.2 Marginal cost

Marginal costs of production, as relevant for the intermediate product market competition, is derived from the following cost function, $C(Y_{jt}) = (1 + T(W(L(Y_{jt}))))W(L(Y_{jt}))L(Y_{jt})$: $mc_{jt} = C'(Y_{jt})$.

Moreover, we have $Y_{jt} = \frac{n_{jt}Y_t}{\gamma m_t}$, and therefore: $\frac{\partial C_{jt}}{\partial n_{jt}} = mc_{jt}\frac{Y_t}{\gamma m_t}$.

D.3 Labor supply elasticity

$$\frac{d\log(L_{jt})}{d\log(W_{jt}\cdot(1+T\left(\frac{W_{jt}}{\bar{W}_t}\right)))} = \frac{dzW_j^{\frac{\beta}{1-\sigma}}}{d(W_{jt}(1+T\left(\frac{W_{jt}}{\bar{W}_t}\right)))}\cdot\frac{W_j\cdot(1+T\left(\frac{W_{jt}}{\bar{W}_t}\right))}{zW_j^{\frac{\beta}{1-\sigma}}} = \frac{\beta(1-\tau)}{1-\sigma} = \varepsilon.$$

D.4 Growth rates

Growth rates of wages are as above in the baseline model, $g_{\bar{W}} = g_w = (g/(g_m g_z))^{\frac{1-\sigma}{\beta}}$. Similarly, for the growth rate of the marginal cost index, g_m, we have:

$$g_m = \frac{mc_j'}{mc_j} = \frac{w_j'[1-\sigma+\beta+(1-\sigma+\beta)T(\frac{w_j'}{\bar{W}'})+(1-\sigma)T'(\frac{w_j'}{\bar{W}'})\frac{w_j'}{\bar{W}'}]}{W_j[1-\sigma+\beta+(1-\sigma+\beta)T(\frac{W_j}{\bar{W}})+(1-\sigma)T'(\frac{W_j}{\bar{W}})\frac{W_j}{\bar{W}}]}$$

Using the g_w result from above:

$$g_m = g_w\frac{[1-\sigma+\beta+(1-\sigma+\beta)T(\frac{g_w W_j}{g_{\bar{W}}\bar{W}})+(1-\sigma)T'(\frac{g_w W_j}{g_{\bar{W}}\bar{W}})\frac{g_w W_j}{g_{\bar{W}}\bar{W}}]}{[1-\sigma+\beta+(1-\sigma+\beta)T(\frac{W_j}{\bar{W}})+(1-\sigma)T'(\frac{W_j}{\bar{W}})\frac{W_j}{\bar{W}}]} = g_w$$

Next, the growth rate of z is defined as follows:

$$g_z = \frac{z'}{z} = \frac{[(g_w^{\frac{\beta}{1-\sigma}}\sum_k W_k^{\frac{\beta}{1-\sigma}})^\sigma g^\beta(\bar{W}Y)^\beta + g_w^{\frac{\beta}{1-\sigma}}\sum_k W_k^{\frac{\beta}{1-\sigma}}]^{-1}}{[(\sum_k W_k^{\frac{\beta}{1-\sigma}})^\sigma(\bar{W}Y)^\beta + \sum_k W_k^{\frac{\beta}{1-\sigma}}]^{-1}}$$

From this a constant g_z requires $g_w^{\frac{\sigma\beta}{1-\sigma}}g^\beta = g_w^{\frac{\beta}{1-\sigma}} \Leftrightarrow g_w = g$.

Then, $g_z = g^{-\frac{\beta}{1-\sigma}}$ and $g = g_Y = g_Q = g_m = g_w = \gamma^X$.

D.5 Set-up, wage

The reference wage is given by:

$$\bar{W} = \int_0^1 W_{o,j(o)}\,do = \int_0^1 \left(\frac{J_H L_H}{L}W_{o,j(o)=h} + \frac{J_L L_L}{L}W_{o,j(o)=l}\right)do \tag{61}$$

Here, $L = J_H L_H + J_L L_L$. Using that $Y_j = \frac{n_j Y}{\gamma m}$, we can rewrite

$$\begin{aligned}\bar{W} &= \frac{J_H L_H W_H + J_L L_L W_L}{L} = \frac{\frac{Y}{\gamma m}\left(J_H \frac{n_H}{s_H} W_H + J_L \frac{n_L}{s_L} W_L\right)}{\frac{Y}{\gamma m}\left(J_H \frac{n_H}{s_H} + J_L \frac{n_L}{s_L}\right)} = \frac{\frac{h}{s_H} W_H + \frac{1-h}{s_L} W_L}{\frac{h}{s_H} + \frac{1-h}{s_L}} \\ &= f_w(h, W_H, W_L)\end{aligned}$$

D.6 Set-up, marginal cost

The cost function is given as:

$$C(Y_j) = \left(1 + \tau\left(\frac{1}{\bar{W}}\left[\frac{Y_j}{s_j z}\right]^{\frac{1-\sigma}{\beta}}\right)\right)\left[\frac{Y_j}{s_j z}\right]^{\frac{1-\sigma}{\beta}} \frac{Y_j}{s_j}$$

which results in the marginal cost as a function of n_j: $mc_j = f_{mc}\left(n_j, s_j, \frac{Y}{mz}, \bar{W}\right)$.

References

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–333.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., & Li, H. (2023). A theory of falling growth and rising rents. *Review of Economic Studies*, *90*(6), 2675–2702.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351. Retrieved August 7, 2024, from http://www.jstor.org/stable/2951599

Akcigit, U., & Ates, S. T. (2023). What Happened to US Business Dynamism? *Journal of Political Economy*, *131*(8), 2059–2124. https://doi.org/10.1086/724289

Azkarate-Askasua, M., & Zerecero, M. (2024). Union and firm labor market power. *Available at SSRN 4323492*.

Bachmann, R., Bayer, C., Stüber, H., & Wellschmied, F. (2024). Monopsony Makes Firms Not Only Small but Also Unproductive: Why East Germany Has Not Converged. *ECONtribute Discussion Paper, No. 328.*

Baqaee, D. R., & Farhi, E. (2019). Productivity and misallocation in general equilibrium*. *The Quarterly Journal of Economics*, *135*(1), 105–163. https://doi.org/10.1093/qje/qjz030

Bassier, I., Dube, A., & Naidu, S. (2022). Monopsony in movers: The elasticity of labor supply to firm wage policies. *Journal of Human Resources*, *57*(S), S50–s86.

Berger, D., Herkenhoff, K., & Mongey, S. (2022). Labor market power. *American Economic Review*, *112*(4), 1147–1193.

Berger, D., Herkenhoff, K., Mongey, S., & Mousavi, N. (2024). Monopsony amplifies distortions from progressive taxes. *AEA Papers and Proceedings*, *114*, 555–60. https://doi.org/10.1257/pandp.20241002

BLS. (2024). Private nonfarm business sector: Total factor productivity [mfpnfbs] [Retrieved from FRED, Federal Reserve Bank of St. Louis].

Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., & Setzler, B. (2023). How much should we trust estimates of firm effects and worker sorting? *Journal of Labor Economics*, *41*(2), 291–322.

Bonhomme, S., Lamadon, T., & Manresa, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, *87*(3), 699–739.

Boppart, T., & Li, H. (2021). Productivity slowdown: Reducing the measure of our ignorance.

Borella, M., De Nardi, M., Pak, M., Russo, N., & Yang, F. (2022). *The importance of modeling income taxes over time. u.s. reforms and outcomes* (Working Paper No. 30725). National Bureau of Economic Research. https://doi.org/10.3386/w30725

Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, *36*.

Cavenaile, L., Celik, M. A., & Tian, X. (2019). Are markups too high? competition, strategic innovation, and industry dynamics. *Competition, Strategic Innovation, and Industry Dynamics (August 1, 2019)*.

Chen, Y., Macaluso, C., & Hershbein, B. (2022). Monopsony in the us labor market. *American Economic Review*, *112*(7), 2099–2138.

Damodaran, A. (2025). R&d statistics by sector (us) [Accessed: March 17, 2025]. https://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/R&D.html

De Ridder, M. (2024). Market power and innovation in the intangible economy. *American Economic Review*, *114*(1), 199–251. https://doi.org/10.1257/aer.20201079

Edmond, C., Midrigan, V., & Xu, D. Y. (2023). How costly are markups? *Journal of Political Economy*, *131*(7), 1619–1675.

Elsby, M. W., Hobijn, B., & Şahin, A. (2013). The decline of the U.S. labor share. *Brookings Papers on Economic Activity*, (FALL 2013), 1–52. https://doi.org/10.1353/eca.2013.0016

Estefan, A., Gerhard, R., Kaboski, J. P., Kondo, I. O., & Qian, W. (2024). *Outsourcing policy and worker outcomes: Causal evidence from a mexican ban* (Working Paper No. 32024). National Bureau of Economic Research. https://doi.org/10.3386/w32024

Fernández-Villaverde, J., Yu, Y., & Zanetti, F. (2025). *Defensive hiring and creative destruction* (tech. rep.). National Bureau of Economic Research.

Garibaldi, P., & Turri, E. (2024). Monopsony in Growth Theory. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4997714

Karabarbounis, L., & Neiman, B. (2014). The Global Decline of the Labor Share*. *The Quarterly Journal of Economics*, *129*(1), 61–103. https://doi.org/10.1093/qje/qjt032

Kirov, I., & Traina, J. (2023). *Labor market power and technological change in us manufacturing* (tech. rep.). Working Paper.

Klenow, P. J., & Li, H. (2025). Entry costs rise with growth. *Journal of Political Economy Macroeconomics*, *3*(1), 43–74. https://doi.org/10.1086/733979

Klette, T. J., & Kortum, S. (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy*.

Kline, P. M. (2025). *Labor market monopsony: Fundamentals and frontiers* (Working Paper No. 33467). National Bureau of Economic Research. https://doi.org/10.3386/w33467

Lamadon, T., Mogstad, M., & Setzler, B. (2022). Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market. *American Economic Review*, *112*(1), 169–212. https://doi.org/10.1257/aer.20190790

Landais, C., Michaillat, P., & Saez, E. (2018). A macroeconomic approach to optimal unemployment insurance: Applications. *American Economic Journal: Economic Policy*, *10*(2), 182–216.

Lehr, N. H. (2024). Does monopsony matter for innovation? *Working Paper*.

Liu, E., Mian, A., & Sufi, A. (2022). Low interest rates, market power, and productivity growth. *Econometrica*, *90*(1), 193–221. https://doi.org/https://doi.org/10.3982/ECTA17408

Macnamara, P., Pidkuyko, M., & Rossi, R. (2024). Marginal tax rates and income in the long run: Evidence from a structural estimation. *Journal of Monetary Economics*, *142*, 103514. https://doi.org/https://doi.org/10.1016/j.jmoneco.2023.09.001

Manning, A. (2021). Monopsony in labor markets: A review. *ILR Review*, *74*(1), 3–26. https://doi.org/10.1177/0019793920922499

McFadden, D. (1977). Modelling the choice of residential location.

Peters, M. (2020). Heterogeneous markups, growth, and endogenous misallocation. *Econometrica*, *88*(5), 2037–2073.

Peters, M., & Walsh, C. (2021). *Population growth and firm dynamics* (NBER Working Paper w29424). National Bureau of Economic Research. https://ssrn.com/abstract=3953951

Rodriguez, F., & Jayadev, A. (2013). The Declining Labor Share of Income. *Journal of Globalization and Development*, *3*(2), 1–18. https://doi.org/10.1515/jgd-2012-0028

Sokolova, A., & Sorensen, T. (2021). Monopsony in labor markets: A meta-analysis. *ILR Review*, *74*(1), pp. 27–55. Retrieved February 6, 2025, from https://www.jstor.org/stable/27111628

Standard & Poor's. (2020). Compustat dataset.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

U.S. Department of Justice and Federal Trade Commission. (2023). Merger guidelines [Accessed: 2025-03-28]. https://www.ftc.gov/system/files/ftc_gov/pdf/2023_merger_guidelines_final_12.18.2023.pdf

Weiss, J. (2023). Market Concentration, Growth, and Acquisitions. *Working Paper*.

Wong, H. C. (2023). *Understanding high-wage firms* (tech. rep.). Mimeo.