A Theory of Falling Growth and Rising Rents

Philippe Aghion, Antonin Bergeaud
Timo Boppart, Peter J. Klenow, Huiyu Li*

October 14, 2022
Click here for the latest version

Abstract

Growth has fallen in the U.S. amid a rise in firm concentration. Market share has shifted to low labor share firms, while within-firm labor shares have actually risen. We propose a theory linking these trends in which the driving force is falling overhead costs of spanning multiple products or a rising efficiency advantage of large firms. In response, the most efficient firms (with higher markups) spread into new product lines, thereby increasing concentration and generating a temporary burst of growth. Eventually, due to greater competition from efficient firms, within-firm markups and incentives to innovate fall. Thus our simple model can generate qualitative patterns in line with the observed trends.

JEL classification: O31, O47, O51.

*Aghion: Collège de France, INSEAD and London School of Economics; Bergeaud: HEC Paris, CEP-LSE and CEPR; Boppart: IIES, Stockholm University and University of St. Gallen; Klenow: Stanford University; Li: Federal Reserve Bank of San Francisco. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve System. We are grateful to Sina Ates, Zsofia Barany, Paco Buera, Basile Grassi, Sam Kortum, Gilles Saint-Paul, and Nicolas Trachter for very helpful discussions. Shelby Buckman, Erin Crust and Lily Seitelman provided excellent research assistance.

1 Introduction

Recent studies have documented the following patterns in the U.S. economy over the past several decades:

1. Slow growth interrupted by a temporary burst of growth
2. Rising firm concentration within industries at the national level
3. Reallocation of market share toward low labor share firms

Figure 1 presents U.S. annual TFP growth in Manufacturing, Trade, and Service industries from the Bureau of Labor Statistics (BLS) KLEMS data.[1] The Figure shows growth accelerating from its 1988–1995 average of 0.8% per year to 2.1% per year from 1996–2005, before falling to just 0.4% per year from 2006–2019.[2]

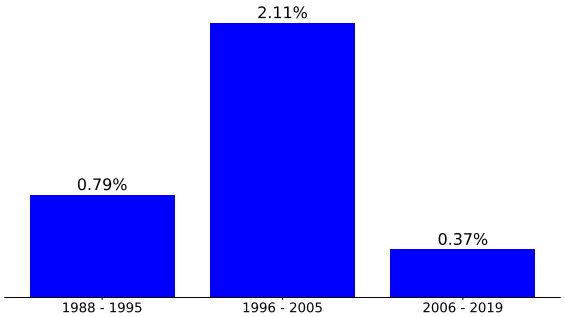
Autor, Dorn, Katz, Patterson and Van Reenen (2020) present evidence on rising concentration within industries, and a tendency of labor share to fall precisely in those industries with bigger increases in concentration. They show that the fall in labor share is due to reallocation of sales toward low labor share firms. Kehrig and Vincent (2021) document the reallocation to low labor share firms in U.S. manufacturing. Similarly, De Loecker, Eeckhout and Unger (2020) and Baqaee and Farhi (2020) report the reallocation of sales shares toward Compustat firms with high ratios of sales to costs of goods sold.[3]

[1] See Figure A1b in the Online Appendix A for U.S. annual TFP growth in all non-farm private industries. The BLS attempts to net out the contribution of both physical and human capital growth to output growth. The BLS sometimes subtracts contributions from R&D and other intellectual property investments; we consistently included this portion in TFP growth as part of what we are trying to explain.

[2] Bergeaud, Cette and Lecat (2016) and Fernald, Hall, Stock and Watson (2017) argue that the recent TFP growth slowdown is statistically significant and predates the Great Recession. Syverson (2017), Aghion, Bergeaud, Boppart, Klenow and Li (2019), and Boppart and Li (2022) contend that the slowdown is unlikely to be fully attributable to growing measurement errors. Aghion et al. (2019) find that measurement error did not increase significantly.

[3] Online Appendix A more fully describes the empirical trends in growth, concentration, and labor share that motivate our modeling effort.

Figure 1: Productivity growth



Source: BLS KLEMS multifactor productivity series. We calculate yearly productivity growth in two digit NAICS manufacturing, trade, and service industries by adding R&D and IP contribution to BLS MFP and then expressing the sum in labor augmenting form. We aggregate industry growth rates using the industry share of labor costs. The figure plots the average productivity growth within each subperiod. The unit is percentage points.

Are these empirical trends linked to one another? Is rising firm concentration slowing growth and driving down the aggregate labor share? Or are large firms more efficient so that their rising market share brings aggregate productivity benefits? These hypotheses are not mutually exclusive, so the answer could be that both are true. If so, there might be a tradeoff between level benefits and adverse longer-run growth effects of rising concentration.

To contribute to this debate, we construct a model of endogenous growth and firm dynamics. There are two sources of firm heterogeneity in our model. The first is product quality, which differs across the product lines of a firm and improves endogenously through creative destruction.[4] The second is process efficiency, which we assume to be common to all product lines of a firm. High process efficiency firms command a higher markup than do low process efficiency firms.

A possible source of persistent heterogeneity in process efficiency across firms is their intangible capital. Firms such as Walmart and Amazon have

[4]Garcia-Macia, Hsieh and Klenow (2019) provide systematic evidence for the importance of creative destruction by incumbents as well as entrants. But Klenow and Li (2021) conclude that innovation by large incumbent firms (not entry and exit) was responsible for the bulk of the speedup and slowdown in growth.

established successful business models and logistics that are evidently hard to copy. Both firms experienced considerable expansion into new geographic and product markets in recent years. Similarly, Amazon and Microsoft have acquired dominant positions in cloud storage and computing due to their logistical advantage over potential competitors. Such firms have achieved a level of process efficiency which is arguably harder to reverse engineer and build upon than quality, which may be more observable.

The story we propose is that overhead costs fell and/or the process efficiency advantage of big firms rose. In our model, these encourage high process efficiency firms to expand into a wider set of product lines. And the expansion of efficient firms into more markets fuels a temporary surge in aggregate productivity growth. Eventually, however, a new innovator on any product line is more likely to face a high-efficiency firm as its competitor. This in turn discourages innovation and growth in the longer term.

Lower overhead costs or a higher process efficiency advantage of larger firms have the additional prediction that within- and between-firm forces move in the opposite direction for markups (or, equivalently, inverse labor shares). High process efficiency firms have higher markups on average across their product lines, so their expansion into more markets pushes the aggregate markup up. Within-firm markups fall, however, as the quality leader on a product line is more likely to face a high process efficiency competitor.

As described, our hypothesis and model can *qualitatively* mimic the broad empirical trends documented in the recent literature. This leaves open the question of whether the forces in our model could play a *quantitatively* important role in these phenomena. To gauge potential magnitudes, we choose parameter values to fit a pre-period (1987–1995) in terms of productivity growth, the level of concentration, the aggregate markup, and the correlation across firms between their labor share and sales share. We then entertain shocks to two parameters to hit two targets. We allow the overhead cost schedule or the process efficiency advantage of the best firms — and the

scale of R&D costs — to change in order to fit the post–2005 values of concentration and productivity growth. We then examine the effect of falling overhead costs or a rising process efficiency edge for big firms on the path of productivity and consumption, and therefore welfare.

Our calibrated fall in overhead costs can explain a nontrivial portion of the "long-run" slowdown in productivity growth: 25 of the 42 basis point fall from before 1995 to after 2006. The rest is explained by a rise in R&D costs in the spirit of Bloom, Jones, Van Reenen and Webb (2020). Alternatively, the rising efficiency advantage of big firms can generate a 63 basis point fall in long-run growth. As that exceeds the 42 basis points decline we observe, it must be paired with *lower* R&D costs to match the decline in growth in the data.

A fall in overhead costs can also explain some of the temporary burst and subsequent slowdown of productivity growth: about 31 basis points of the 1.31 percentage point acceleration, and 56 basis points of the 1.74 percentage point slowdown. As noted, growth ends up 25 basis points lower after 2005 than before 1996 in response to lower overhead costs. We find that the long-run decline in growth outweighs the short-run burst of growth, leaving consumption-equivalent welfare about 12 basis points lower.

A widening productivity lead for large firms can similarly explain some of the growth burst and slowdown: about 32 basis points of the 1.31 percentage point acceleration, and 95 basis points of the 1.74 percentage point slowdown. The growth drag in the long run is 63 basis points, as mentioned. Taking into account the full transition path, we find that a process efficiency gain for the most efficient firms actually lowers welfare by a substantial -3.28 percent.

Our calibrated model's implications might seem paradoxical: how could lower overhead costs for all firms or higher process efficiency for a subset of firms lower welfare? This outcome underscores the second-best nature of the decentralized equilibrium in our model. Our model illustrates the possibility that the rise of superstar firms, even if it yields short-run productivity gains, might result in lower long-run growth and welfare.

Like us, Hsieh and Rossi-Hansberg (2022) model how shocks (in their case to fixed and marginal costs) can lead to rising national concentration within industries and a burst of process improvements. They present evidence for Wholesale Trade, Retail Trade, and Services on the expansion of large firms into more geographic markets. They do not model long-run growth or changes in labor share due to markup dispersion. We view the multi-product firms in our model as in the spirit of their multi-location firms, with our fixed innovation costs playing the role of fixed costs of adding locations.

Our paper is complementary to a number of other recent studies of falling growth and rising concentration. These studies feature different driving forces than ours. In Akcigit and Ates (2022) the driving force is declining imitation rates, in Peters and Walsh (2022) it is declining population growth, and in Liu, Mian and Sufi (2022) it is declining interest rates. Our paper also relates to Hopenhayn, Neira and Singhania (2022) and Chatterjee and Eyigungor (2022), who study rising concentration, and to recent papers on forces behind the declining aggregate labor share such as Karabarbounis and Neiman (2014), Farhi and Gourio (2018), Kaymak and Schott (2020), Barkai (2020), Koh, Santaeulàlia-Llopis and Zheng (2020) and Eggertsson, Robbins and Wold (2021).

Perhaps closest to our paper is De Ridder (2021). As we do, De Ridder (2021) models changes in technology that allow large firms to expand, which in turn lowers long-run growth after an initial burst of growth. The mechanism behind falling growth in his model revolves around convex R&D costs. In contrast, we center on a general equilibrium channel whereby the expansion of superstar firms to new product lines lowers the expected markup for any potential innovator. As a result, our model distinctively generates a decline in within-firm markups (and a rise in within-firm labor shares) along with falling growth despite rising aggregate rents.

The rest of the paper proceeds as follows. Section 2 lays out our model. Section 3 solves for the balanced growth path analytically and performs some

qualitative comparative statics. Section 4 calibrates the model to see how much a drop in overhead costs or a rise in the efficiency edge of large firms can contribute to the secular trends in growth and labor share. Section 5 solves numerically for the transition dynamics in the model, and explores how much it can account for the burst and then slowdown of growth. Section 6 concludes.

2 A model of innovation with heterogeneous firms

This section lays out an endogenous growth model with firm heterogeneity. In later sections we will use it to draw out predictions for concentration, the distribution of markups, and growth.

2.1 Preferences

Time is discrete and the economy is populated by a representative household who chooses a path of consumption C and wealth A to maximize

$$U_0 = \sum_{t=0}^{\infty} \beta^t \log\left(C_t\right),$$

subject to $A_{t+1} = (1 + r_t)A_t + w_t L - C_t$, a standard no-Ponzi game condition, and initial wealth $A_0 > 0$. Here r is the real interest rate, w is the real wage, and L is the endowment of labor, which is inelastically supplied to the labor market.

The usual Euler equation resulting from household optimization is given by

$$\frac{C_{t+1}}{C_t} = \beta(1 + r_{t+1}).$$

2.2 Production of final output

In each period, a final output good is produced competitively using a continuum of intermediate inputs according to a Cobb-Douglas technology[5]

$$Y = \exp\left(\int_0^1 \log\left[q(i)y(i)\right]di\right).$$

Here $y(i)$ denotes the quantity and $q(i)$ the quality of product i. This structure yields demand for each product i as

$$y(i) = \frac{YP}{p(i)}, \tag{1}$$

where the aggregate price index (which we normalize to 1 in each period) is

$$P \equiv \exp\left(\int_0^1 \log\left[p(i)/q(i)\right]di\right).$$

2.3 Production and market structure for intermediate inputs

There are J firms indexed by j. J is "large" such that firms take P as given. Each firm j has the knowledge to produce quality $q(i,j) \geq 0$ in a specific market $i \in [0,1]$. There are two sources of heterogeneity across firms: (i) product-specific quality $q(i,j)$ which evolves endogenously with innovation; and (ii) permanent heterogeneity in firm-specific process efficiency.

We denote firm-specific process efficiency by $\varphi(j)$. A firm with process efficiency $\varphi(j)$ can produce in any line i with the linear technology

$$y(i,j) = \varphi(j) \cdot l(i,j), \tag{2}$$

where $l(i,j)$ is labor used by firm j to produce output $y(i,j)$ in product line i. We assume the heterogeneity in process efficiency is permanent. This heterogeneity in process efficiency will translate into persistent differences in

[5] When there is no ambiguity we suppress time subscripts.

markups and labor share across firms. The linear technology in (2) applies irrespective of the specific quality $q(i,j)$ at which firm j produces in line i.

We explain below how product-specific quality changes endogenously due to innovations. For the static firm problem within a period, the line-specific quality of a firm $q(i,j)$ can be taken as given. And because labor is fully mobile and firms take wages as given, all firms face the same wage rate in equilibrium. Hence, the marginal cost of firm j in line i is $w/\varphi(j)$.

2.4 Pricing

In each market i firms engage in Bertrand competition. This implies that only the firm with the highest quality-adjusted productivity $q(i,j) \cdot \varphi(j)$ will in equilibrium be active in a given market. We denote the leading firm in line i by $j(i)$ and the second-highest quality producer by $j'(i)$. Hence the quality-adjusted productivity of the leader in line i is $q(i,j(i)) \cdot \varphi(j(i))$, whereas it is $q(i,j'(i)) \cdot \varphi(j'(i))$ for the second-best firm. Under Bertrand competition, the price setting of the leading firm is constrained by the second-best producer. The leader will set its quality-adjusted price equal to the quality-adjusted marginal cost of the second-best firm. Formally, we then have

$$\frac{p(i,j(i),j'(i))}{q(i,j(i))} = \frac{w}{\varphi(j'(i)) \cdot q(i,j'(i))}.$$

Note that the equilibrium price in line i depends on the process efficiency of the *second-best* firm as well as the quality difference between the competing firms.

The markup in line i, the price of a unit divided by the marginal cost, is

$$\mu(i,j(i),j'(i)) \equiv \frac{p(i,j(i),j'(i))}{w/\varphi(j(i))} = \frac{q(i,j(i)) \cdot \varphi(j(i))}{q(i,j'(i)) \cdot \varphi(j'(i))}.$$

The markup is increasing in the quality gap $q(i,j(i))/q(i,j'(i))$ and the process efficiency gap $\varphi(j(i))/\varphi(j'(i))$ between the leading and the second-best firm. Profits of the leader in line i are $Y\,[1 - 1/\mu(i,j(i),j'(i))]$. This follows from the

demand function (1) with P normalized to one. Hereafter, we express profits in a line i relative to aggregate output Y and denote these scaled profits by $\Pi(i) \equiv 1 - 1/\mu(i, j(i), j'(i))$.

2.5 Innovation and productivity growth

The quality distribution evolves endogenously over time as a result of innovations. Any firm j can engage in R&D to acquire a patent to produce a product at higher than existing quality. More specifically, by investing $x_t(j) \cdot \psi_r \cdot Y_t$ units of final output in R&D in period t, $x_t(j)$ product lines are randomly drawn among the lines in which firm j is currently not actively producing. In these randomly drawn lines the highest existing quality is multiplied by a factor $\gamma > 1$ and the innovating firm j obtains a perpetual patent to produce at this higher quality level from the next period $t+1$ onward.

As the total number of lines is normalized to 1, the aggregate rate of creative destruction is then given by[6]

$$z_{t+1} = \sum_{j=1}^{J} x_t(j).$$

That is, for any given line, an innovation arrives in $t+1$ with probability z_{t+1}. These quality improvements are the source of long-run growth.

2.6 Boundary of the firm

Given the constant R&D cost of acquiring a line through innovation and the fact that firms with higher process efficiency make higher expected profits $\Pi(i)$ in an additional line, more productive firms have a stronger incentive to invest in R&D. To prevent the firm with the highest productivity from taking over all lines, we assume that firms have to pay a per-period overhead cost which is

[6] We assume that no two innovations can arrive on the same line in a given period. With continuous time this would be fulfilled automatically.

a convex function of the number of markets they span. More specifically, we assume a quadratic per-period overhead cost

$$\frac{1}{2}\psi_o\, n(j)^2\, Y, \tag{3}$$

with $\psi_o > 0$, where $n(j)$ denotes the number of lines in which firm j owns the highest quality patent. We assume a quadratic cost to obtain an analytical solution to the firm's dynamic problem. The convexity of the overhead cost in $n(j)$ gives rise to a natural boundary of the firm. High process efficiency firms will operate more lines than low process efficiency firms, but no firm type will operate all lines.[7]

It may be helpful to compare our model to Klette and Kortum (2004), a benchmark model in the firm dynamics and growth literature. Klette and Kortum assume a convex cost of acquiring extra product lines through creative destruction, and a non-diminishing value of adding lines (the firm's value function is linear in n along a balanced growth path). By contrast, we assume a linear cost of innovating on a new line and convex overhead costs.[8] Hence our model allows us to do comparative statics with respect to the scalar ψ_o that changes a firm's span of control without altering the R&D technology. Another difference with Klette and Kortum is that we assume that each firm operates a continuum of lines, so that there is no firm exit in our model.

2.7 Markups with binary process efficiency levels

For simplicity we assume in the following two types of firms. A fraction ϕ of all firms has high process efficiency φ_H whereas the remaining fraction $1-\phi$ has a low process efficiency φ_L. We denote their efficiency ratio by $\Delta \equiv \varphi_H/\varphi_L > 1$.

[7]Convex overhead costs in (3) would seem to give firms an incentive to divide in order to reduce their combined overhead costs. We assume such divestment does not occur because of a combination of transaction costs, asymmetric information, and difficulty in maintaining process efficiency.

[8]Our model shares some features with Luttmer (2011), in which more efficient firms endogenously expand into more products because their efficiency extends across product lines.

We further assume $\gamma > \Delta$ so that the firm with the highest quality is the active producer irrespective of its type or the type of the second-best firm.[9]

Given the two process efficiency levels (high and low) there are four potential cases for markups $\mu(i)$ and profits $\Pi(i)$ in a given line i (the latter expressed before overhead costs and relative to aggregate output) :[10]

1. A high productivity leader $\varphi(j(i)) = \varphi_H$ facing a high productivity second-best firm $\varphi(j'(i)) = \varphi_H$ in line i, giving rise to

$$\mu(i) = \gamma \text{ and } \Pi(i) = 1 - \frac{1}{\gamma}.$$

2. A high productivity leader $\varphi(j(i)) = \varphi_H$ facing a low productivity second-best firm $\varphi(j'(i)) = \varphi_L$ in line i, giving rise to

$$\mu(i) = \Delta\gamma \text{ and } \Pi(i) = 1 - \frac{1}{\Delta\gamma}.$$

3. A low productivity leader $\varphi(j(i)) = \varphi_L$ facing a high productivity second-best firm $\varphi(j'(i)) = \varphi_H$ in line i, giving rise to

$$\mu(i) = \frac{\gamma}{\Delta} \text{ and } \Pi(i) = 1 - \frac{\Delta}{\gamma}.$$

4. A low productivity leader $\varphi(j(i)) = \varphi_L$ facing a low productivity second-best firm $\varphi(j'(i)) = \varphi_L$ in line i, giving rise to

$$\mu(i) = \gamma \text{ and } \Pi(i) = 1 - \frac{1}{\gamma}.$$

[9]The existing literature has so far emphasized the corner case where Δ is equal to one, thus less than γ. If $\Delta > \gamma$, in general, the highest quality producer no longer coincides with the leading (active) firm. It then matters whether new innovations build on the currently produced product or on the highest quality product. With the former assumption no interior balanced growth path will exist, i.e., we will have either no long-run growth or the high productivity firms take over the whole market. With the latter assumption, additional combinations will arise, e.g., where a low productivity firm that is two steps ahead competes with a high productivity producer.

[10]To ease notation we denote the markup in line i, $\mu(i, j(i), j'(i))$, simply by $\mu(i)$.

2.8 Labor shares

Our baseline model abstracts from physical capital; labor is the only factor of production. Furthermore, both R&D expenditure and overhead costs are denominated in final output and are treated as investments rather than intermediate inputs. These last two assumptions are made to avoid a mechanical effect of overhead and R&D spending on the aggregate labor share. Hence in our baseline framework the aggregate labor share is just equal to one minus the share of profits before overhead and R&D costs.

Because of the Cobb-Douglas technology in final good production, revenue for each product is equal to Y. Labor costs in a line i equal $wl(i) = Y/\mu(i)$. Integrating both sides over i yields $wL = Y \int_0^1 \mu(i)^{-1} di$. Dividing $wl(i)$ by wL, the wage bill (or employment) share of product line i is

$$\frac{l(i)}{L} = \frac{1}{\mu(i)} \frac{1}{\int_0^1 \frac{1}{\mu(\iota)} d\iota}.$$

The employment share on a line, $l(i)/L$, is inversely proportional to the markup on the line. This comes from revenue being equalized across lines.

The aggregate labor share λ is then given by the inverse of the cost-weighted markup:

$$\lambda \equiv \frac{wL}{Y} = \frac{1}{\int_0^1 \mu(i) l(i)/L \, di} = \int_0^1 \mu(i)^{-1} \, di.$$

The aggregate labor share depends non-trivially upon the full distribution of markups across lines. This distribution is determined by the types of the leader and second-best firm across lines.

Consider a firm j with $n(j)$ lines that faces a fraction $h(j)$ of high-type second-best firms and a remaining fraction $1 - h(j)$ of low-productivity second-best firms. If firm j is itself of high type, its overall labor share is given by

$$\lambda_H(h(j)) = h(j)\frac{1}{\gamma} + (1 - h(j))\frac{1}{\gamma\Delta}. \tag{4}$$

In contrast, if firm j is low type its overall labor share is given by

$$\lambda_L(h(j)) = h(j)\frac{\Delta}{\gamma} + (1 - h(j))\frac{1}{\gamma}. \tag{5}$$

Faced with the same share of high-type competitors $h(j)$, high-productivity firms have a lower labor share as they can charge on average higher markups. Hence the model generates persistent differences in labor shares across firms.[11] As the composition of competitors $h(j)$ is endogenous, the model can generate changes in the labor share within firms over time.

2.9 Dynamic firm problem

There are two individual state variables in the problem of a firm of type k: the number of lines firm j operates, $n(j)$, and the fraction of high productivity second-best producers, $h(j)$, the firm faces in its lines. Each firm then chooses how many new lines to innovate upon, $x(j) \geq 0$, to maximize the net present value of its flow of profits net of all costs. Denoting per-period profits net of labor and overhead costs (and expressed relative to aggregate output Y) for high and low productivity type firms by π_H and π_L, we have:

$$\pi_H(n(j), h(j)) = n(j) - \frac{n(j)h(j)}{\gamma} - \frac{n(j)\,(1 - h(j))}{\gamma\Delta} - \frac{1}{2}\psi_o n(j)^2, \tag{6}$$

and

$$\pi_L(n(j), h(j)) = n(j) - \frac{n(j)h(j)\Delta}{\gamma} - \frac{n(j)\,(1 - h(j))}{\gamma} - \frac{1}{2}\psi_o n(j)^2. \tag{7}$$

These profits (expressed relative to Y) only depend on the firm's type $k = H, L$ and the individual states $n(j)$ and $h(j)$, and are otherwise time invariant. Hereafter we will refer to profit functions (6) and (7) as "firm profits" and the partial derivative of these functions with respect to n as "marginal firm profits." Letting S denote the aggregate fraction of lines operated by high-productivity

[11] See Hsieh and Klenow (2009) and David and Venkateswaran (2019) for evidence of persistent differences in revenue labor productivity across firms.

firms, the problem of a firm of type $k = H, L$ can be written as

$$V_{0,k}(n_0, h_0) = \max_{\{x_t, n_{t+1}, h_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} Y_t \left[\pi_k(n_t, h_t) - x_t \psi_r\right] \prod_{s=0}^{t} \left(\frac{1}{1+r_s}\right)$$

subject to

$$n_{t+1} = n_t(1 - z_{t+1}) + x_t, \tag{8}$$

$$h_{t+1} n_{t+1} = h_t n_t (1 - z_{t+1}) + S_t x_t. \tag{9}$$

As firm profits differ by type k, there are two firm value functions indexed by $k = H, L$. The constraint (9) captures that the firm understands that a newly added line is drawn from the pool of lines with a fraction S_t of high types. Hence by choosing the innovation rate the firm can influence next period's number of lines in which the firm faces a high type competitor h_{t+1}.

2.10 Market clearing and resource constraints

We close the model with the following market clearing conditions. As they have to hold each period we suppress time indices. First, final output will be used for consumption C, total overhead costs O, and total R&D expenditures Z:

$$Y = C + O + Z, \tag{10}$$

where

$$O = \sum_{j=1}^{J} \frac{1}{2} \psi_o n(j)^2 Y \quad \text{and} \quad Z = \sum_{j=1}^{J} x(j) \psi_r Y.$$

There is no free entry and the number of firms is fixed. Hence total profits from selling at a markup over marginal cost may exceed the total investments in R&D and overhead costs. We call such net profits "rents" R. As total firm sales are Y and total firm costs consist of overhead, R&D, and labor, we have $R = Y - O - Z - wL$.

Labor is used as a variable input by the producers of different intermediate

product lines. Labor and asset market clearing conditions imply

$$L = \sum_{j=1}^{J} \int_0^1 l(j,i)\, di \quad \text{and} \quad \sum_{j=1}^{J} V(j) = A,$$

where $l(j,i)$ denotes labor used by firm j on line i.

In addition, we have the equations defining the aggregate share of lines operated by high types and an accounting equation that states that all lines are operated by some firm:[12]

$$S = \sum_{j=1}^{\phi J} n(j) \;\text{ and }\; 1 = \sum_{j=1}^{J} n(j). \tag{11}$$

Finally, there is an equation relating aggregate output to the distribution of process efficiency, quality levels, and markups

$$Y = Q \frac{\varphi_L \Delta^S \exp\left[-\int_0^1 \log\left(\mu(i)\right) di\right]}{\int_0^1 \left(\mu(i)\right)^{-1} di} L. \tag{12}$$

Here $Q = \exp\left[\int_0^1 \log\left(q(i,j)\right) di\right]$ denotes the geometric average quality level and Δ^S is the parameter Δ raised to the power S.

An equilibrium in this economy is a path of allocations and prices that jointly solve the household and firm problems and is consistent with the market clearing and accounting equations stated above.

Since output is a function of the full distribution of markups across product lines, the equilibrium path is a function of the initial joint distribution of product lines $n(j)$ and level of competition $h(j)$ across firms. We assume that all firms of the same type $k = H, L$ start out with the same level of n_0 and h_0.[13] Using the law of large numbers, firms of the same type will then be identical along the entire equilibrium path. Therefore, only two firm problems — one

[12] Here we assume that the high productivity type firms are indexed by $j = 1, 2, \ldots, \phi J$.

[13] Along a balanced growth path, which we will study next, such symmetry within types is automatically ensured.

for a high type and one for a low type — need to be solved. The aggregate state vector can then be summarized by S and the shares of high second-best firms h_H and h_L in lines operated by high- and low-productivity firms.

With the two "representative" firms, aggregate labor productivity can be expressed in terms of these aggregate state variables (S, h_L, h_H) and the level of average quality Q as

$$\frac{Y}{L} = Q \cdot \varphi_L \Delta^S \cdot \frac{\Delta^{(1-S)h_L - S(1-h_H)}}{Sh_H + (1-S)(1-h_L) + S(1-h_H)\frac{1}{\Delta} + (1-S)h_L\Delta}.$$

Aggregate labor productivity is the product of three terms. The first term Q captures the geometric average level of quality across product lines. The second term, $\varphi_L \Delta^S$, captures the aggregate level of process efficiency. If $S = 0$ then aggregate process efficiency is just the level of the low type φ_L, whereas if $S = 1$ aggregate process efficiency is equal to the high level $\varphi_H = \varphi_L \Delta$. The third and final term, which we call allocative efficiency, captures the output distortion due to markup dispersion. If $S = h_H = 1$ or $S = h_L = 0$ this final term is equal to 1 (no dispersion of markups since markups are equal to γ in all lines). In all other cases the third term is smaller than one.

3 Solving for the balanced growth path

Here we solve the balanced growth path (BGP) analytically, and analyze various comparative static effects on outcomes in the long-run (i.e., on the new BGP).

3.1 Balanced growth path definition

We define a balanced growth path in the following way:

Definition 1 *A balanced growth path is an equilibrium path along which all quantities and prices change at constant rates.*

Note that this definition just restricts growth rates to be constant but allows them to differ across variables and also to take a value of zero. Together with the resource constraint (10) and the Euler equation the definition implies that the interest rate, the rate of creative destruction z, and S all have to be constant along a BGP. In the following we denote these levels by $r^\star$, $z^\star$ and $S^\star$, respectively.[14] As the BGP definition also implies a stationary markup distribution, it follows from (12) that the growth rates of Y and average quality Q coincide, i.e.,

$$\frac{Y_{t+1}}{Y_t} = \frac{Q_{t+1}}{Q_t} = \gamma^{z^\star} = 1 + g^\star. \tag{13}$$

In the following we restrict our attention to a BGP with interior $z^\star \in (0,1)$ and $S^\star \in (0,1)$. The following assumption ensures the existence of such an interior BGP:

Assumption 1 *We assume the following parameter restrictions hold:*

$$\frac{\Delta - 1}{\gamma} < \frac{\psi_o}{\phi J}, \tag{14}$$

and

$$0 < \frac{1}{\psi_r} - \frac{1-\beta}{\beta} - \frac{1}{\psi_r} \frac{\frac{\psi_o}{J} + \frac{1}{\gamma}}{1 - (1-\phi)\phi \frac{(\Delta-1)^2}{\gamma\Delta} \frac{J}{\psi_o}} < 1. \tag{15}$$

Restriction (14) in Assumption 1 ensures that the low type firms are active on the BGP with $S^\star < 1$.[15] This is fulfilled as long as neither the productivity differential Δ nor the number of high-productivity firms ϕJ is too large. Restriction (15) ensures a positive but less than certain rate of creative destruction, $0 < z^\star < 1$. It is fulfilled as long as ψ_r relative to β is neither too small nor too large.

Note that, as Assumption 1 ensures $z^\star > 0$, it implicitly ensures strictly positive output growth $g^\star > 0$ given (13). With strictly positive growth and

[14] The constancy of the interest rate directly follows from the Euler equation and is given by $r^\star = \frac{1+g^\star}{\beta} - 1$, where $g^\star$ is the net growth rate of consumption (and output).

[15] With $\frac{\Delta-1}{\gamma} \geq \frac{\psi_o}{\phi J}$ there exists a trivial BGP with $n_L^\star = 0$, $n_H^\star = 1/(\phi J)$, $S^\star = 1$, and $z^\star = (1 - 1/\gamma - \psi_o/(\phi J))/\psi_r + 1 - 1/\beta$, where $0 < (1 - 1/\gamma - \psi_o/(\phi J))/\psi_r + 1 - 1/\beta < 1$ needs to be imposed to ensure that the high-type firms invest strictly positive amounts and that the rate of creative destruction is less than 100%, i.e., $z^\star \in (0,1)$.

creative destruction, there will be symmetry within types along a BGP meaning that all high-productivity firms have $n_H^\star$ lines and all low-productivity firms have $n_L^\star$ lines along such a path. Furthermore, for the number of lines of all firms to be constant over time, a firm's R&D activities must be proportional to the firm's number of products. That is, $x(j)^\star = n(j)^\star z^\star$, $\forall j$. Finally, because all firms draw new lines from a stationary distribution, along a BGP they all face the same share of high-productivity second-best firms in their lines:

$$h(j)^\star = S^\star \; \forall j. \tag{16}$$

We next solve for the interior balanced growth path.

3.2 Characterization of the interior balanced growth path

With $h(j)^\star = S^\star$, equations (6) and (7) imply that per-period firm profits for high- and low-type firms (relative to total output) are,

$$\pi_H(n, S^\star) = n\left(1 - \frac{S^\star}{\gamma} - \frac{1 - S^\star}{\gamma\Delta}\right) - \frac{1}{2}\psi_o n^2, \tag{17}$$

and

$$\pi_L(n, S^\star) = n\left(1 - \frac{S^\star\Delta}{\gamma} - \frac{1 - S^\star}{\gamma}\right) - \frac{1}{2}\psi_o n^2. \tag{18}$$

Denote v as the value of a firm relative to total output, $v \equiv V/Y$. The number of products per firm n becomes the only individual state variable in the firm problem along a BGP, so we can write $v_k = v_k(n)$, $k = H, L$. High- and low-productivity firms then solve the following Bellman equations:

$$v_H(n) = \max_{n' \geq n(1-z^\star)} \{\pi_H(n, S^\star) - (n' - n(1 - z^\star))\psi_r + \beta v_H(n')\}, \tag{19}$$

$$v_L(n) = \max_{n' \geq n(1-z^\star)} \{\pi_L(n, S^\star) - (n' - n(1 - z^\star))\psi_r + \beta v_L(n')\}. \tag{20}$$

We denote their solutions as $n' = f_H(n)$ and $n' = f_L(n)$.

Along the BGP, the two accounting equations in (11) become

$$S^\star = n_H^\star \phi J$$

$$n_H^\star \phi J + n_L^\star (1-\phi) J = 1.$$

Finally, we must have

$$n_H^\star = f_H(n_H^\star) \quad \text{and} \quad n_L^\star = f_L(n_L^\star).$$

These equations fully characterize the BGP. The two dynamic programming problems (19) and (20) are very simple since π_H and π_L are quadratic functions of n from (17) and (18).

By imposing Assumption 1 we ensure the existence of an interior BGP solution wherein $S^\star \in (0,1)$ and $z^\star \in (0,1)$. Along this interior BGP, the policy and value functions can be characterized in closed form.[16]

The next two propositions characterize the interior BGP solution and prove that Assumption 1 is sufficient for the existence of such a BGP.

Proposition 1 *An interior balanced growth path is characterized by a quadruple* $(n_H^\star, n_L^\star, S^\star, z^\star)$ *that fulfills*

$$\phi J n_H^\star = S^\star \tag{21}$$

$$(1-\phi) J n_L^\star + \phi J n_H^\star = 1, \tag{22}$$

as well as the following research optimality conditions for high- and

[16] Let us denote the expected marginal BGP profits per line before overhead (and expressed relative to Y) by $\widetilde{\pi}_H = 1 - S^\star/\gamma - (1-S^\star)/(\Delta\gamma)$ and $\widetilde{\pi}_L = 1 - \Delta S^\star/\gamma - (1-S^\star)/\gamma$. Then, for any $n \leq \bar{n}_k/(1-z^\star)$, where $\bar{n}_k \equiv (\widetilde{\pi}_k + (1-z^\star)\psi_r - \psi_r/\beta)/\psi_o$, we have the policy function $f_k(n) = \bar{n}_k$ and the value function $v_k(n) = \widetilde{\pi}_k n - \frac{1}{2}\psi_o n^2 - \psi_r(\bar{n}_k - (1-z^\star)n) + \beta(\widetilde{\pi}_k \bar{n}_k - \frac{1}{2}\psi_o \bar{n}_k^2 - \psi_r z^\star \bar{n}_k)/(1-\beta)$, for $k = H, L$. See Online Appendix B.1 for details.

low-productivity firms:

$$\psi_r = \frac{1 - S^\star/\gamma - (1 - S^\star)/(\gamma\Delta) - \psi_o n_H^\star}{1/\beta - 1 + z^\star} \tag{23}$$

$$\psi_r = \frac{1 - S^\star\Delta/\gamma - (1 - S^\star)/\gamma - \psi_o n_L^\star}{1/\beta - 1 + z^\star}. \tag{24}$$

Proof. By definition $S^\star \in (0, 1)$ along an interior BGP. This implies that $n_k^\star$ and $x_k^\star$ are positive for $k = H, L$. Thus, both firm policy functions satisfy the first-order condition for the Bellman equation. For the high type this is

$$\psi_r = \beta \frac{\partial v_H(n')}{\partial n'}.$$

Using the envelope theorem we have

$$\frac{\partial v_H(n')}{\partial n'} = 1 - S^\star/\gamma - (1 - S^\star)/(\gamma\Delta) - \psi_o n' + (1 - z^\star)\psi_r.$$

Using the fact that $n' = n_H^\star$ along a BGP then yields the research optimality condition of the high-type firm. The research optimality condition of the low type firm is derived in an analogous way. ■

The intuition for the two research optimality conditions is straightforward. On the BGP the marginal cost of innovating in a line (ψ_r) equals the marginal (expected) value of having an additional line. For the high-type firm, this marginal value is equal to the expected marginal period profit before overhead costs of a line served by a high-type producer: $1 - S^\star/\gamma - (1 - S^\star)/(\gamma\Delta)$ minus the marginal overhead cost $\psi_o n_H^\star$. These terms are divided by $1/\beta - 1 + z^\star$ (due to time discounting and the probability $z^\star$ of losing the additional line in each future period) to arrive at the marginal value of a product line for a high-type firm. The research optimality condition of the low-type looks analogous and simply differs by the firm profit function.

Equation (21)–(24) are four equations in the four unknowns $(n_H^\star, n_L^\star, S^\star, z^\star)$.

We solve them explicitly and use them to show that Assumption 1 guarantees an interior solution, as well as to solve for all the other endogenous variables.

Proposition 2 *An interior balanced growth path has the following properties:*

(i) The share of lines operated by high-productivity firms is equal to

$$S^\star = \frac{\frac{1}{1-\phi} + \frac{\Delta-1}{\gamma\Delta}\frac{J}{\psi_o}}{\frac{1}{(1-\phi)\phi} - \frac{(\Delta-1)^2}{\gamma\Delta}\frac{J}{\psi_o}}, \tag{25}$$

and the rate of creative destruction is given by

$$z^\star = \frac{1}{\psi_r} - \frac{1-\beta}{\beta} - \frac{1}{\psi_r}\frac{\frac{\psi_o}{J} + \frac{1}{\gamma}}{1-(1-\phi)\phi\frac{(\Delta-1)^2}{\gamma\Delta}\frac{J}{\psi_o}}. \tag{26}$$

(ii) High-productivity firms operate more lines than low-productivity firms:

$$n_H^\star > n_L^\star.$$

(iii) The labor share of a high-type firm is given by

$$\lambda_H^\star = S^\star\frac{1}{\gamma} + (1-S^\star)\frac{1}{\gamma\Delta}, \tag{27}$$

which is strictly smaller than the labor share of a low type firm

$$\lambda_L^\star = S^\star\frac{\Delta}{\gamma} + (1-S^\star)\frac{1}{\gamma}. \tag{28}$$

Finally, the aggregate labor share is given by

$$\lambda^\star = S^\star\lambda_H^\star + (1-S^\star)\lambda_L^\star. \tag{29}$$

Proof. An interior balanced growth path is characterized by the four equations in Proposition 1. Replacing $n_H^\star$ and $n_L^\star$ in (23) and (24) by $S^\star/(\phi J)$ and $(1-S^\star)/(J(1-\phi))$, respectively, and solving the two equations for $S^\star$ and $z^\star$

yields the unique solution in part (i). Note that restriction (14) ensures $S^\star < 1$ and restriction (15) ensures $0 < z^\star < 1$. Finally, note that $S^\star > 0$ is always guaranteed since (14) implies $\frac{\psi_o}{\phi J} > \frac{\Delta-1}{\gamma} > \frac{\Delta-1}{\gamma}\frac{\Delta-1}{\Delta}(1-\phi)$, as $\frac{\Delta-1}{\Delta}(1-\phi) < 1$. Assumption 1 is sufficient to ensure the existence of an interior BGP.

For part (ii), by combining (23) and (24) the difference in the number of products can be expressed as

$$n_H^\star - n_L^\star = \frac{S^\star(\Delta-1)^2}{\gamma\Delta\psi_o} + \frac{\Delta-1}{\gamma\Delta\psi_o} > 0.$$

The labor shares follow from (4), (5) and (16). This proves part (iii). ■

The intuition for (ii) in Proposition 2 is that high process efficiency firms charge higher average markups. Consequently their incentive to undertake R&D is higher and they push up into a steeper section of the convex overhead cost schedule. That is, they operate more lines than low efficiency firms do along the BGP. A corollary is that we have $S^\star > \phi$ since high-productivity firms are larger (in terms of sales per firm) than low-productivity firms. High- and low-productivity firms also differ in their employment, but the employment difference is smaller than the sales difference because high-productivity firms charge higher markups — see part (iii) of Proposition 2.

On the BGP, $S^\star$ can be viewed as a measure of market concentration, whereas $z^\star$ pins down the long-run growth rate of the economy. Note that the endogenous values $S^\star$ and the rate of creative destruction $z^\star$ only depend on the ratio ψ_o/J and not on the individual level of ψ_o or J.

3.3 Comparative static effects along the balanced growth path

As the theory admits closed-form expressions of all main outcome variables of interest along the balanced growth path (see Proposition 2), the framework lends itself to characterizing comparative static effects analytically. Table 1 summarizes the signs of the effects of changes in ψ_o, Δ, γ, and ψ_r on

concentration, within- and between-firm labor shares, and the growth rate. Here we focus on the intuition behind these effects, relegating their derivations to Online Appendix B.2. In the Online Appendix we also state the conditions for the different signs in cases where the comparative static effect cannot be signed unambiguously (denoted by a "$\pm$" in Table 1).

The sign of the effects on $S^\star$ are all unambiguous, i.e., do not depend on parameter constellations. Hence, our theory makes clear predictions about long-run concentration. The resulting changes in $S^\star$, however, give rise to opposing effects on the within- and between-firm labor share. To see this, note that (27) and (28) imply that the labor share for both firm types is increasing in $S^\star$. The intuition is that higher $S^\star$ implies that firms — irrespective of their type — are more likely to face a high-type second-best firm and can therefore expect to charge lower markups on average the higher is $S^\star$. The aggregate labor share, however, is a weighted average of the labor share of the two types as shown in equation (29). As $\lambda_H^\star = \lambda_L^\star/\Delta < \lambda_L^\star$, an increase in $S^\star$ gives rise to a negative between-firm effect on the aggregate labor share. Higher $S^\star$ means a higher sales share for higher-than-average markup firms. It is therefore at the core of our framework that changes in $S^\star$ have opposing within- and between-firm effects, so that the sign of the overall effect on the aggregate labor share depends on parameter conditions.

We now discuss the comparative static effects in Table 1 in detail, starting with the effects of a change ψ_o. Along an interior balanced growth path, expected marginal firm profits have to be equalized across the two types of firms (see (23) and (24)). For a given $S^\star$ (i.e., for a given expected markup) an increase in ψ_o decreases expected marginal firm profits, incentivizing both types of firms to innovate less until they end up in a flatter section of the overhead cost schedule. However, in equilibrium, the number of products cannot decrease for both types. Hence, $n_H^\star$ has to fall and $n_L^\star$ has to rise, decreasing concentration $S^\star$ and increasing markups and expected marginal firm profits for each type of firm. The lower $S^\star$ has (as explained above)

Table 1: Effect of changes in parameter values along the BGP

Effect of an increase in:	$S^\star$	$\lambda_H^\star$	$\lambda_L^\star$	$\lambda^\star$	$z^\star$	$g^\star$
ψ_o	$-$	$-$	$-$	$\pm$	$\pm$	$\pm$
Δ	$+$	$\pm$	$+$	$\pm$	$-$	$-$
γ	$-$	$-$	$-$	$\pm$	$+$	$+$
ψ_r	0	0	0	0	$-$	$-$

Notes: Each entry displays the sign of the partial derivative of the column variable with respect to the row parameter. "$\pm$" means the sign is ambiguous. See Online Appendix B.2 for derivations.

opposing within- and between-firms effects on the aggregate labor share, leaving the effect on its sign ambiguous. (We show in Online Appendix B.2 that the between firm effect dominates and the aggregate labor share $\lambda^\star$ decreases in ψ_o if and only if $S^\star > 1/2$.) Finally, there are also two opposing effects of an increase in ψ_o on growth: a higher marginal overhead cost decreases the incentive to invest in R&D, whereas the increase in within-firm markups stimulates it. As a consequence, the sign of the overall effect on the growth rate depends on parameter values.

Higher Δ raises the productivity edge of high-type firms relative to low type firms. As a consequence, the high-type firms will expand until they reach a steeper section of the overhead cost schedule, such that $S^\star$ increases. The increase in $S^\star$ is a force that pushes down markups of both firm types. But the increase in the productivity gap is a force reducing markups for low type firms while increasing markups for high-type firms. Thus, the labor share unambiguously increases for low type firms, whereas the effect on the labor share of high-type firms is ambiguous. The between-firm effect adds another force that decreases the aggregate labor share leaving this effect ambiguous. Finally, the drop in within firm markups for low types leads to a fall in growth

as Δ increases.

An increase in γ increases the innovation step size and therefore directly leads to more growth $g^\star$ (for a given $z^\star$). In addition, a rising γ increases the general markup firms can charge irrespective of their type (and both $\lambda_H^\star$ and $\lambda_L^\star$ fall). As the relative difference in markups between high- and low-type firms falls this lowers concentration $S^\star$. This in turn decreases within-firm labor shares, whereas the compositional effect adds a positive force on the labor share. As a consequence, the overall effect on the labor share depends on parameter conditions. Finally, the generally higher within firm markups increase the incentive to undertake R&D activity and therefore increase the rate of creative destruction, $z^\star$.

In contrast to changes in ψ_o, Δ, and γ, a change in ψ_r has no compositional effect and leaves concentration $S^\star$ unaffected. As ψ_r increases the only effect is that R&D productivity falls for all firms and growth declines as a result.

To recap, our theory can generate a productivity slowdown, rising concentration, and rising (falling) within (between) firm labor shares in response to either a drop in ψ_o or an increase in Δ. In Section 4 below we will gauge the quantitative size of these effects in a simple calibration. As changes in ψ_r are orthogonal to changes that work through concentration $S^\star$, we also change ψ_r values to fit the residual effect on the growth rate.

A decrease in γ could give rise to similar qualitative effects as a fall in ψ_o or a rise in Δ. In addition to compositional effects, however, a decrease in γ has a direct negative effect on within-firm markups and growth. Furthermore, a decrease in γ implies that the observed trends have to be accompanied by shrinking as opposed to rising rents. We therefore focus in the quantitative section on changes in ψ_o and Δ to account for rising concentration.

3.4 Theoretical extensions

We kept our baseline model parsimonious to show the minimum ingredients needed to speak to the empirical facts described in Section 1 and in Online Appendix A. This tractable model can be augmented in various ways to explore the same mechanisms in richer environments. In Online Appendix C, we consider two extensions of the model. We replace the Cobb-Douglas aggregation across goods with a CES final good aggregator in Section C.1. This introduces a finite monopoly markup and hence leads to less markup dispersion. We also generalize from two to an arbitrary number of firm types in terms of their permanent process efficiency in Section C.2.

4 Quantitative comparative statics on the BGP

We now investigate how a decline in overhead costs or an increase in the process efficiency gap might contribute quantitatively to a slowdown in long-run productivity growth. We define the initial BGP period as 1987–1995 and the ending BGP period as 2005–2019. We calibrate six parameters $(\phi, \psi_o, \psi_r, \Delta, \beta, \gamma)$ to match six moments in the initial BGP. And we infer changes in two of the parameters — ψ_o and ψ_r or Δ and ψ_r — to match the concentration level and productivity growth rate in the ending BGP. We then calculate changes in growth on the BGP when only the overhead cost parameter ψ_o or the efficiency gap parameter Δ changes.

4.1 Initial balanced growth path

The six parameters we have to calibrate are: 1) the percent of high-type firms ϕ; 2) the initial overhead cost parameter ψ_o^0; 3) the initial R&D cost parameter ψ_r^0; 4) the initial process efficiency gap $\Delta^0 > 1$ between high- and low-type firms; 5) the quality step size $\gamma > 1$; and 6) the discount factor $0 < \beta < 1$.

The six moments we match are: 1) the percentile (share) of firms

represented by the top 20 firms within a given industry; 2) the share of sales going to the largest 20 firms within industries in 1987 from Autor et al. (2020); 3) the semi-elasticity of firm labor shares with respect to firm sales within four-digit industries divided by the average labor share in Manufacturing, Trade, and Services from the BLS KLEMS dataset;[17] 4) average markups in Manufacturing, Trade, and Service industries as estimated by Hall (2018) over 1988–2015;[18] 5) the average annual rate of productivity growth over 1987–1995 in Manufacturing, Trade, and Services from the BLS KLEMS dataset;[19] and 6) the real interest rate from Farhi and Gourio (2018) for 1980–1995.

Table 2 and Table 3 display the targets and the calibrated parameter values, respectively. We set the share of high-type firms ϕ to 0.374% to match the fraction of top 20 firms.[20] Concentration $S^\star$ is monotonically decreasing in the overhead cost parameter ψ_o in the model. Hence the level of concentration helps to pin down ψ_o^0 to $1.1 \cdot 10^{-3}$. Next, the semi-elasticity of labor shares with respect to size becomes more negative with higher Δ, helping to set Δ^0 to 1.183. That is, the high-type firms enjoy about 18% higher process efficiency. Given Δ and $S^\star$, the average markup increases with the quality step γ. To match the average markup in the data, the model asks for a 25% increase in

[17]We aggregate concentration and the semi-elasticity from Autor et al. (2020) for Manufacturing, Retail Trade, Wholesale Trade, and Services using value of production weights from the BLS KLEMS dataset. The semi-elasticity is based on sales as opposed to value-added, as the underlying quinquennial Census data lacks information on intermediate inputs outside manufacturing. In our model labor shares and markups are inversely related to each other, so we target this cross-sectional pattern in labor share to discipline how markups vary with firm size in our quantitative model.

[18]To calculate the price/cost markup for these sectors, we aggregate industry level Lerner indices (one minus the inverse of markups) estimated by Hall (2018) using sector output shares provided by Hall (2018) and then convert the resulting aggregate index to markups.

[19]We follow growth accounting standards and calculate empirical growth rates using the log difference of aggregate productivity in adjacent periods. Hence when we calibrate we map $z^\star \log(\gamma)$ in the model to the empirical growth rate. For the same reason, we report $g_{t+1} = z_{t+1} \log(\gamma)$ in the transition dynamics section and compare the acceleration and deceleration of g_{t+1} to the speedup and slowdown in the data.

[20]Autor et al. (2020) report the average sales shares of the top 20 firms and the average number of firms within 4-digit industries in the Manufacturing, Trade, and Service sectors. We use this information to infer that the top 20 firms are the top 0.374% of firms on average.

quality upon innovation (γ = 1.246).[21] All else equal, the growth rate in the model decreases with ψ_r, which scales the cost of R&D. We obtain ψ_r^0 = 1.460. Finally, for a given growth rate of the economy, the real interest rate decreases with the discount factor β. Matching the real interest rate requires β = 0.950. As shown in Table 2, the model has the same number of parameters as moments and can fit all of them.

Table 2: Baseline calibration targets

Targeted	Years	Data	Model
1. percentile of top 20 firms	1987	0.374	0.374
2. sales share of top 20 firms	1987	45.9	45.9
3. elasticity of labor share wrt sales	1987	-3.07	-3.07
4. price/cost markup	1988–2015	1.24	1.24
5. productivity growth	1987–1995	0.79	0.79
6. real interest rate	1980–1995	6.10	6.10

Sources: 1, 2 and 3 come from Autor et al. (2020); 4 from Hall (2018); 5 from BLS KLEMS series; and 6 from Farhi and Gourio (2018).

Table 3: Baseline parameter values

Calibrated	Parameter	Value
1. percent of H-type firms	ϕ	0.374
2. overhead costs	ψ_o^0	$1.1 \cdot 10^{-3}$
3. productivity gap	Δ^0	1.183
4. quality step	γ	1.246
5. R&D costs	ψ_r^0	1.460
6. discount factor	β	0.950

[21]Note that our markup target is not as high as the inverse of labor's share in the data. Although our model has no capital input, we think of labor in our model as a stand-in for both capital and labor inputs.

4.2 Balanced growth path after parameters change

Table 4 displays the moments in the new BGP when ψ_o and ψ_r change to match the "long-run" empirical changes in concentration and growth from pre-1996 to post-2005. As displayed in Table 1, concentration declines with ψ_o and does not change with ψ_r. Therefore, the rise in concentration pins down the change in ψ_o. We find that the overhead cost parameter ψ_o must decline by 19% to generate the rise in concentration seen in the data. As shown in Table 1, the effect of ψ_o on growth is non-monotonic. At the calibrated parameter values we find that the decline in ψ_o leads to lower long-run growth but does not generate the entire growth decline observed in the data. As a consequence, the data ask for the R&D cost parameter ψ_r to increase by about 11%. This is reminiscent of Bloom, Jones, Van Reenen and Webb (2020), who argue that growth is held down by ever-rising research costs.

Since both ψ_o and ψ_r contribute to the decline in long-run growth, in Table 5 we isolate the contribution of ψ_o alone. This contribution can be calculated in two ways: 1) the change in growth when only ψ_o changes, and 2) the change in growth when ψ_o does not change relative to when both ψ_o and ψ_r change. These two methods differ because the growth rate is a nonlinear function of the parameter ψ_o, in particular in a way that depends on the value of ψ_r. The average of the two says the decline in ψ_o contributed 25 out of the 42 basis point decline in long-run productivity growth.

Table 4: Calibrated change in ψ_o and ψ_r to fit the ending BGP

	Change	Targeted change	Data	Model
1. overhead costs ψ_o	-19.4%	concentration	12.1	12.1
2. R&D costs ψ_r	10.5%	productivity growth	-0.42	-0.42

Source: 1: Autor et al. (2020), change in the sales share of the top 0.374% firms between 1987 and 2012. 2: BLS KLEMS. Columns 'Data' and 'Model' are in percentage points.

Alternatively, the process efficiency advantage of high-type firms, Δ, may have increased over time. In our model an increase in Δ can lower the

Table 5: Contribution of overhead costs to the decline in long-run growth

	1. ψ_o, ψ_r	2. only ψ_r	1. minus 2.	3. only ψ_o	ψ_o contribution
change in growth	-0.42	-0.19	-0.24	-0.26	**-0.25**

Note: Each column displays the percentage point change in the long-run growth rate when the parameters in the column header changes as in Table 4. Column 1. matches the change in the data. Column 'ψ_o contribution' equals the average of columns 3. and 4.

long-run growth rate by raising market concentration and decreasing within-firm markups (see Table 1). Table 6 displays the moments in the new BGP when we keep ψ_o constant and vary relative process efficiency Δ and ψ_r to match the empirical changes in concentration and growth. The model asks Δ to increase by 4% and ψ_r to *decline* by 11%. Thus, a rise in Δ alone could explain more than 100% of the decline in long-run growth. As shown in Table 7, the decline in growth from the change in Δ alone exceeds the growth decline in the data, 63 vs. 42 basis points.[22]

Table 6: Calibrated change in Δ and ψ_r to fit the ending BGP

	Change	Targeted change	Data	Model
1. efficiency gap Δ	4.0%	concentration	12.1	12.1
2. R&D costs ψ_r	-11.2%	productivity growth	-0.42	-0.42

Source: 1: Autor et al. (2020), change in the sales share of the top 0.374% firms between 1987 and 2012. 2: BLS KLEMS. Columns 'Data' and 'Model' are in percentage points.

Table 8 displays the value of selected endogenous variables in the initial and ending balanced growth paths. Lower overhead costs or a higher efficiency gap increase the share of products and sales at the high efficiency firms (i.e., they increase $S^\star$).[23] With the rise in $S^\star$, within-firm markups *decline* for both firm types because the next best producer is more likely to be a

[22] In a similar spirit, De Ridder (2021) obtains negative effects of rising concentration on growth due to a disadvantageous R&D technology of large firms. Adding this mechanism would amplify the negative effect of concentration on growth. See Online Appendix D.2 for a discussion.

[23] Employment concentration increases by less than sales concentration because the high efficiency firms charge higher markups ($\mu_H > \mu_L$).

Table 7: Contribution of efficiency gap to the decline in long-run growth

	1. Δ, ψ_r	2. only ψ_r	1. minus 2.	3. only Δ	Δ contribution
change in growth	-0.42	0.25	-0.67	-0.60	**-0.63**

Note: Each column displays the percentage point change in the long-run growth rate when the parameters in the column header changes as in Table 6. Column 1. matches the change in the data. Column 'Δ contribution' equals the average of columns 3. and 4.

Table 8: Initial vs. ending balanced growth paths

	$z^\star$	$S^\star$	μ	μ_H	μ_L	r	Z/Y	O/Y	R/Y
Initial	3.6	45.9	1.24	1.36	1.15	6.1	5.3	3.1	10.8
Ending (ψ_o, ψ_r)	1.7	58.0	1.24	1.33	1.13	5.7	2.7	4.0	12.5
Ending (Δ, ψ_r)	1.7	58.0	1.23	1.35	1.10	5.7	2.2	5.0	11.8

Note: "Ending (ψ_o, ψ_r)" is the new BGP when ψ_o and ψ_r change as in Table 4. "Ending (Δ, ψ_r)" is the new BGP when Δ and ψ_r change as in Table 6. μ, μ_H and μ_L denote the aggregate and firm-level markups. Aggregate labor share equals $1/\mu$. Aggregate labor income, total R&D expenditure Z, total overhead cost O and rents sum up to GDP Y. Creative destruction rate ($z^\star$), concentration ($S^\star$), interest rate ($r^\star$), and shares of GDP are in percent.

high-type firm. Despite the decline in within-firm markups, the aggregate markup changes little because of the rising market share of high-type firms. That is, the between effect roughly cancels out the within effect.

The decline in within-firm markups discourages innovation by both firm types, lowering the rate of creative destruction and growth.[24] R&D spending as a share of total output declines (Z/Y falls).[25] Meanwhile, the rise in concentration leads to a rise in overhead costs as a share of output (higher O/Y) even in the case where ψ_o declines. The decline in the R&D share exceeds the rise in the overhead cost share. This, combined with a stable aggregate labor share (the inverse of the aggregate markup), implies an increase in the share of rents in GDP by about 1 to 1.7 percentage points (higher R/Y). Finally,

[24] A firm's marginal profit declines by -0.017 (13.5%) when ψ_o declines by the amount shown in Table 4. The increase in $S^\star$ lowers expected firm profits before overhead costs in an additional line for all firms. It also raises the marginal overhead cost for the high efficiency firms because the overhead cost schedule is convex.

[25] Note that Trade and Services report little R&D, so this is probably not captured well by R&D data, which is predominantly in manufacturing and software.

the 42 basis point decline in the growth rate generates a decline in the real interest rate from 6.1% in the initial BGP to 5.7% in the ending BGP. This is in the direction of the decline estimated by Farhi and Gourio (2018) from 6.1% for 1980–1995 to 4.5% from 2000–2016. The changes in the level of allocative efficiency and process efficiency when ψ_o declines are relatively small. We display this in Figure A4 of Online Appendix D.[26]

Theoretically, the effects of ψ_o and Δ on the aggregate labor share may be non-monotonic (see Table 1). Table 9 says that, at the calibrated parameter values, lowering the overhead cost ψ_o to match the rise in concentration generates more than 100% of the within and almost 30% of the between changes in labor share as reported in Autor et al. (2020) for Manufacturing, Trade, and Services. Varying Δ to match the change in concentration generates almost triple the within changes seen in the data, and about 60% of the between changes. Most of the within change comes from an increase in the labor share of the low type because the within change is the change in the unweighted average of firm labor share $\phi\lambda_H + (1-\phi)\lambda_L$ and most of the firms are low type ($1-\phi = 99.626\%$). For both changes in Δ and ψ_o, the within and between components roughly cancel out, leaving the aggregate labor share relatively stable in our model.

Table 9: Change in labor share in the long run

	Total	Within	Between
Data change over 1987–2012	-5.97	1.65	-7.62
Model (vary ψ_o)	-0.01	2.19	-2.20
Model (vary Δ)	0.34	4.89	-4.55

Note: The Data row displays changes in labor share components reported in Autor et al. (2020). We aggregate Manufacturing, Trade and Services using their shares of BLS KLEMS production in 1987. The Model rows show the difference in labor share components between the initial and new Balanced Growth Paths (BGPs) as a percent of the aggregate labor share in the initial BGP. All entries are in percentage points.

[26]Process efficiency generally decreases with ψ_o while allocative efficiency decreases with ψ_o when $S < 0.5$ and increases with ψ_o when $S > 0.5$.

Table 9 shows within changes in labor share as a *percent* of the initial aggregate labor share. In levels, within labor shares rise 1.8 and 4.0 percentage points in response to the the calibrated changes in ψ_o and Δ, respectively. These increases are intimately tied to the BGP growth declines generated by the changes in ψ_o and Δ, as higher labor shares within firms are equivalent to lower markups and hence lower marginal return to R&D. More precisely, the research optimality conditions (23) and (24) together imply

$$z^\star = \frac{1 - [\phi\lambda_H + (1-\phi)\lambda_L]}{\psi_r} - \frac{\psi_o}{\psi_r J} - \frac{1-\beta}{\beta}. \tag{30}$$

This expression says that, all else equal, the BGP rate of creative destruction $z^\star$ declines as the labor share within firms increases (higher $\phi\lambda_H + (1-\phi)\lambda_L$). Substituting ψ_r^0 and the level of within changes corresponding to Table 9 into expression (30) yields a $z^\star$ decline of 1.2 and 2.7 percentage points and a corresponding $g^\star$ decline of 0.27 and 0.60 percentage points. Hence the rise in within firm labor shares accounts for all of the decline in growth caused by a ψ_o decline or a Δ increase (see column 3. of Tables 5 and 7). Furthermore, the growth impact of Δ is twice as big as the growth impact of ψ_o because the within increase in firm labor shares is twice as big for Δ.[27]

5 Transition dynamics

Our analysis so far has been based on comparative statics across balanced growth paths. We now consider how changes in overhead costs and in the process efficiency gap might contribute to the ten-year burst in U.S. productivity growth from 1995 to 2005. We compute our model's transition in response to the parameter changes in Table 4 and 6 to see the potential contribution of ψ_o and Δ to the acceleration and deceleration of growth.[28]

[27]The decline in ψ_o has a direct positive effect on R&D in (30), but this is quantitatively negligible for the calibrated values of ψ_o^0 and ψ_o^1.

[28]See Online Appendix E for a description of the computation method.

5.1 Transition dynamics from declining overhead costs

After a drop in ψ_o, our model generates a surge in productivity growth along the transition. The reason for this temporary burst is twofold: 1) the general equilibrium force that reduces the incentive to innovate — stiffer competition as S_t increases — is only realized over time. On impact, as ψ_o decreases, the incentive to do R&D increases and quality growth increases; and 2) the new BGP with a higher $S^\star$ exhibits higher average process efficiency as the efficient firms operate a larger fraction of the product lines. This static efficiency gain is realized along the transition and contributes to the burst of growth.

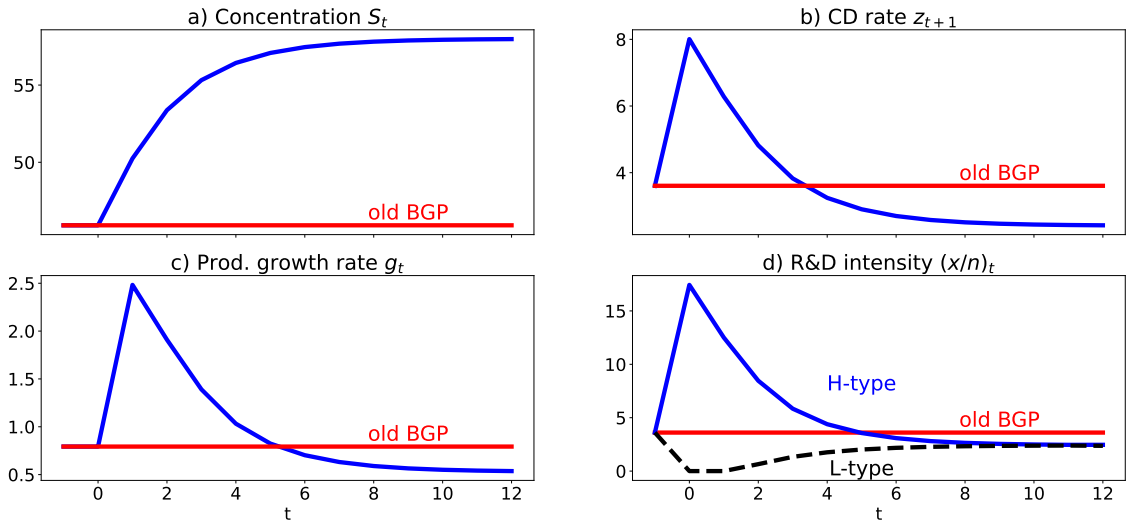
Figure 2 depicts the transition dynamics for the share of lines operated by the high-type firms (S_t), the rate of creative destruction (z_{t+1}), and productivity growth after the overhead cost parameter ψ_o declines in year 0. Concentration S_t rises sharply and converges to the new BGP after around 10 years. On impact, there is a spike in the rate of creative destruction (z_{t+1}) that generates higher productivity growth than the initial BGP for about 5 years. In addition to the rise in innovation, aggregate process efficiency rises by about 2% as more products are produced by the high productivity firms. Allocative efficiency does not change significantly.

The connection between rising concentration and (at least initially) rising productivity growth is qualitatively consistent with the findings of Autor et al. (2020) that industries with increasing concentration saw faster TFP growth over 5-year periods. Similarly, Covarrubias, Gutiérrez and Philippon (2020) find rising concentration in the 1990s initially went along with rising productivity growth across U.S. industries, whereas later in the 2000s it was associated with slower productivity growth.

The rise in creative destruction comes from a jump in innovation by high-type firms, which in turn lifts concentration over time. Figure 2d shows the innovation rate (x/n) by firm type. The innovation rate of low-type firms actually falls. Eventually the innovation rate of both types converges to a level

below the one on the initial BGP.[29]

Figure 2: Transition dynamics when ψ_o declines to match rise in concentration



Note: The blue line plots the transition dynamics when overhead cost ψ_o declines as in Table 4 in period 0 while other parameters stay the same. The unit is percent.

Table 10 displays the contribution of the decline in ψ_o alone to the ten-year burst in growth. We find that the decline in overhead costs might account for 24% of the rise in productivity growth from 1995–2005 and 32% of the subsequent decline in productivity growth.

The drop in ψ_o raises consumption growth in the short run but reduces consumption growth in the long run. Figure 3 compares the path of consumption following the reduction in ψ_o with the initial BGP. Following the decline in ψ_o, consumption drops sharply in the first period as high-efficiency firms increase R&D and overhead investments. Consumption then recovers and is above the initial balanced growth path for about 25 years. Eventually the long-run slowdown in innovation and growth takes its toll and consumption falls below its old balanced growth trajectory.

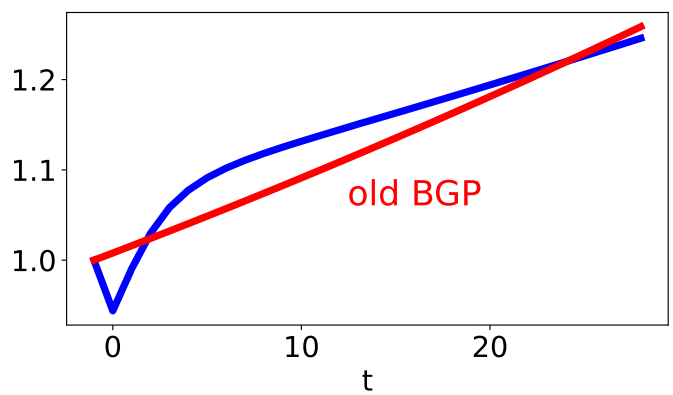
[29] If one makes the strong assumption that new product lines are associated with plant entry, then this behavior qualitatively matches the pattern in Figure A2b in the Online Appendix, wherein only the largest firms experienced a burst of plant entry rate during the high growth period from 1995–2005.

Table 10: Contribution of decline in overhead costs to growth burst and decline

	Data	1. ψ_o, ψ_r	2. ψ_r	3. 1. minus 2.	4. ψ_o	5. ψ_o contribution
Acceleration	1.31	0.16	-0.17	0.33	0.30	**0.31 (24.0%)**
Deceleration	1.74	0.59	0.02	0.57	0.56	**0.56 (32.4%)**

Note: For data, 'Acceleration' = growth over 1995–2005 minus growth over 1987–1995. 'Deceleration' = growth over 1995–2005 minus growth over 2005–2018. In the model, 'Acceleration' = growth over period 0 to 9 minus initial BGP growth. 'Deceleration' = growth over period 0 to 9 minus new BGP growth. Each column displays the acceleration and deceleration in growth when the parameters in the column header change as in Table 4. Column 'ψ_o contribution' equals the average of columns 3. and 4. Entries are in percentage points.

Figure 3: Transition dynamics for consumption when ψ_o declines



Note: The blue line plots consumption relative to period -1. Overhead cost ψ_o declines in period 0 as in Table 4, while other parameters stay the same. The red line plots the old BGP.

Given the consumption response in Figure 3, it is natural to ask whether welfare is higher or lower because of the drop in ψ_o. Recall that utility from a consumption path is given by

$$U_0 = \sum_{t=0}^{\infty} \beta^t \log C_t \equiv U(\{C_t\}_{t=0}^{\infty}).$$

The change in welfare can therefore be evaluated in (permanent) consumption-equivalent terms, ξ, using

$$U\left(\{(1+\xi)C_t^{old}\}_{t=0}^{\infty}\right) = \frac{\log(1+\xi)}{1-\beta} + U(\{C_t^{old}\}_{t=0}^{\infty}) = U(\{C_t^{new}\}_{t=0}^{\infty}),$$

where $\{C_t^{new}\}_{t=0}^{\infty}$ and $\{C_t^{old}\}_{t=0}^{\infty}$ are paths of consumption with and without a change in ψ_o and/or ψ_r. When we allow both ψ_o to fall and ψ_r to rise, welfare is *lower* by the same amount as a permanent 2.6% decrease in consumption. As shown in Table 11, however, when only ψ_o declines, welfare declines by only 0.35%. When only ψ_r rises, welfare declines even more (by 2.8%). Averaging these two changes, the decline in ψ_o lowers welfare by the same amount as a modest but permanent 0.12% decline in consumption. Even though the decline in overhead costs generates a burst of growth through a combination of a permanent boost in process efficiency and a temporary surge in innovation, overall it lowers welfare by reducing long-run innovation.[30]

Table 11: Contribution of the decline in overhead costs to welfare

	1. ψ_o and ψ_r	2. only ψ_r	3. 1. minus 2.	4. only ψ_o	5. ψ_o contribution
ξ%	-2.64	-2.75	0.11	-0.35	**-0.12%**

Note: Each column displays the welfare change in consumption-equivalent percentage terms when the parameters in the column header change to the value in Table 4. Column 5, the ψ_o contribution, equals the average of columns 3 and 4.

5.2 Transition dynamics from higher efficiency gap

Figure 4 displays the path of concentration, innovation, productivity growth, and firm-level R&D intensity as the economy transitions to a new BGP after Δ increases as in Table 6 in period 0 while other parameters stay the same. The transition dynamics from an increase in the efficiency gap Δ are qualitatively similar to the dynamics following a decline in overhead costs. One notable

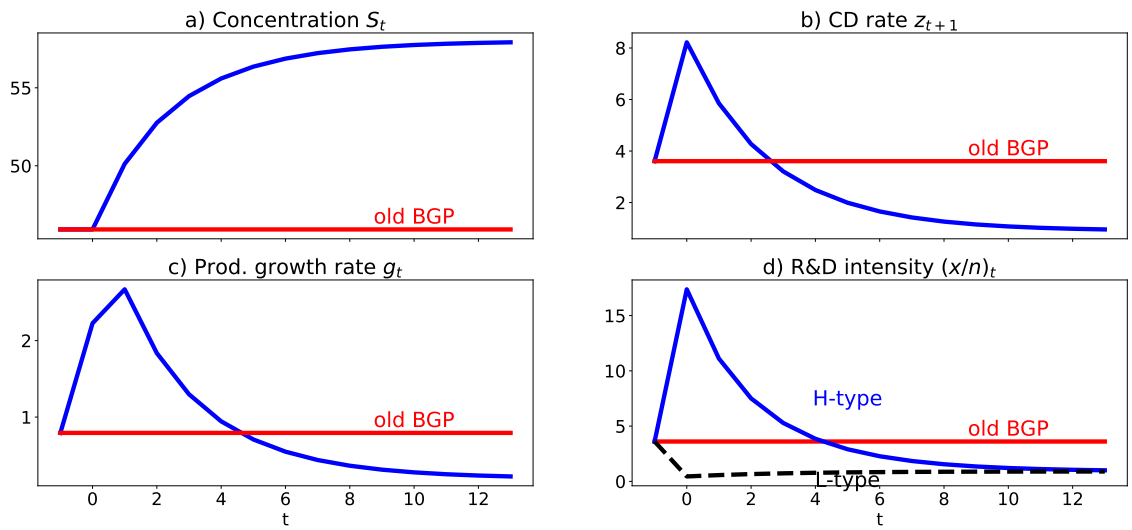
[30] In our framework, the optimal rate of creative destruction is given by

$$z^P = \frac{1}{\psi_r} - \frac{1/\beta - 1}{\log(\gamma)}\left(1 + \log(\Delta)\frac{1}{2}J\phi(1-\phi)(1/\beta - 1)\frac{\psi_r}{\psi_o}\frac{\log(\Delta)}{\log(\gamma)}\right) - \frac{\psi_o}{\psi_r}\frac{1}{2J}.$$

When we evaluate the optimal growth rate at the baseline combination of parameters and decrease ψ_o by 19.4% as in Table 4 (but hold ψ_r unchanged) we see that the optimal growth rate decreases mildly from 9.70% to 9.69%, which is much smaller than the decline in the equilibrium growth rate (1 basis point here versus the 25 basis point decline in Table 5).

difference is the behavior of productivity growth in Figure 4c in period 0. The decline in ψ_o does not affect productivity growth in period 0 because innovation and concentration are determined by firm decisions in the previous period. In contrast, an increase in Δ has an immediate effect on productivity growth because it directly increases the efficiency of high-type firms (φ_H increases) in period 0. This is captured by the Δ^S term in equation (12).

Figure 4: Transition dynamics when Δ rises to match rise in concentration



Note: The blue line plots the transition dynamics when efficiency gap Δ changes as in Table 6 in period 0 while other parameters stay the same. The unit is percent.

Table 12 displays the contribution of the change in Δ to the burst and decline of productivity growth from 1996–2005. Similar to the decline in overhead costs, the change in Δ can generate about one-quarter of the acceleration in growth rate. The rise in Δ can account for one-half of the deceleration in growth, however, larger than the roughly one-third contribution of falling overhead costs that we reported in Table 10.

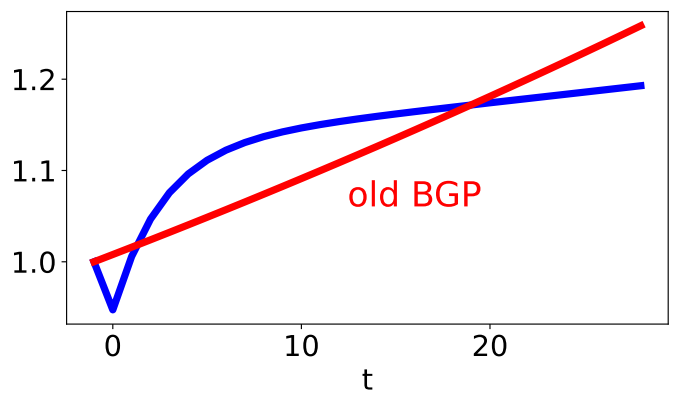
Figure 5 compares the path of consumption after an increase in Δ relative to the initial BGP. As in the case of a decline in overhead costs, consumption first drops relative to the initial BGP as high-efficiency firms invest more in R&D. Consumption remains above the initial BGP for about two decades as the

Table 12: Contribution of increase in efficiency gap to growth burst and decline

	Data	1. Δ, ψ_r	2. ψ_r	3. 1. minus 2.	4. Δ	5. Δ contribution
Acceleration	1.31	0.52	0.22	0.30	0.34	**0.32 (24.4%)**
Deceleration	1.74	0.94	-0.02	0.97	0.94	**0.95 (54.8%)**

Note: For data, 'Acceleration' = growth over 1995–2005 minus growth over 1987–1995. 'Deceleration' = growth over 1995–2005 minus growth over 2005–2018. In the model, 'Acceleration' = growth over period 0 to 9 minus initial BGP growth. 'Deceleration' = growth over period 0 to 9 minus new BGP growth. Each column displays the acceleration and deceleration in growth when the parameters in the column header change as in Table 6. Column 'Δ contribution' equals the average of columns 3. and 4. Entries are in percentage points.

Figure 5: Transition dynamics for consumption when Δ increases



Note: The blue line plots consumption relative to period -1 when efficiency gap Δ changes as in Table 6 in period 0 while other parameters stay the same.

economy reaps the benefit of higher process efficiency and higher productivity growth. Compared to the decline in overhead costs, consumption drops below the initial BGP earlier, in about twenty years instead of twenty-five years. As a result, the welfare loss is larger than the decline in ψ_o. Table 13 reports that the rise in Δ lowers welfare by the same amount as a permanent 3.3% decline in consumption.

Table 13: Contribution of the increase in efficiency gap to welfare

	1. Δ and ψ_r	2. only ψ_r	3. 1. minus 2.	4. only Δ	5. Δ contribution
ξ%	0.04	4.00	-3.96	-2.60	**-3.28%**

Note: Each column displays the welfare change in consumption-equivalent percentage terms when the parameters in the column header change to the value in Table 6. Column 5, the Δ contribution, equals the average of columns 3 and 4.

6 Conclusion

We developed a model of innovation-led growth with intrinsic firm heterogeneity. We solved analytically for the balanced growth path and its qualitative properties. We then solved for transition dynamics numerically to gauge whether a fall in overhead costs or an increase in the productivity edge of large firms could account for a portion of the U.S. growth experience over the past 30 years, in particular (i) a productivity growth slowdown (after a temporary burst in growth); (ii) rising concentration at the national level; and (iii) opposing between- and within-firm changes in labor's share of income.

According to our model, lower overhead costs and higher productivity at leading firms enticed such firms to expand their product boundaries. Because the most efficient firms enjoyed higher markups, when they expanded their reach into more markets this pushed down the aggregate labor share. And it led to a temporary surge of productivity growth. Within-firm markups eventually fell for both high and low productivity firms, as any given firm became more likely to face high productivity competitors. This force ultimately reduced within-firm markups and dragged down innovation and growth. Despite the temporary burst of growth, welfare fell in our calculation (either modestly or more significantly, depending on whether overhead costs fell or the productivity edge of leading firms widened).

Future work could try to document whether, how much, and why overhead costs fell or productivity gaps widened. Information Technologies (IT) —

linked to computer hardware and software — may have played a role. Crouzet and Eberly (2019) and Lashkari, Bauer and Boussard (2021) document that bigger firms invest a higher share of their sales in intangibles and IT, respectively. The former evidence is for U.S. firms and the latter for French firms. Bessen (2020) finds that sales grew faster for the largest firms in IT-intensive industries. Babina, Fedyk, He and Hodson (2021) demonstrate that larger firms invested more in Artificial Intelligence in the last decade, and moved into more markets as a result. Jiang (2021) links IT access and investments to firm expansion into new geographic markets. This features prominently in Hsieh and Rossi-Hansberg (2022) as well.

We emphasize creative destruction in our model, which contributes about one-fourth of overall growth according to Garcia-Macia et al. (2019) and Klenow and Li (2021). Our model could be extended to include other sources of growth, such as new variety creation. A decline in span-of-control costs could alter the incentive for large versus small firms to introduce new varieties, and in turn the profitability of future creative destruction on those varieties.

Our baseline framework is based on leapfrogging innovations. Our model does not feature a positive *escape from competition effect* as in Akcigit and Ates (2022) or Liu, Mian and Sufi (2022). One could introduce such an effect into our model by allowing for step-by-step innovation. Such a channel could allow creative destruction to affect the incentive of incumbents to improve their own products, another importance source of growth.

One could explore optimal tax and subsidy policies in our quantitative framework. The decentralized equilibrium is suboptimal due to markup dispersion across products as well as knowledge spillovers and business stealing across firms. Falling overhead costs and higher productivity at the best firms may increase welfare in the presence of optimal R&D subsidies.

Finally, our framework is well suited for how analyzing how competition policy — such as mergers and acquisitions, data access, or firm breakups — affect growth. We leave these extensions of our analysis for future research.

References

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li, "Missing growth from creative destruction," *American Economic Review*, 2019, *109* (8), 2795–2822.

Akcigit, Ufuk and Sina T Ates, "What happened to U.S. business dynamism?," *Journal of Political Economy*, 2022.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Babina, Tania, Anastassia Fedyk, Alex Xi He, and James Hodson, "Artificial intelligence, firm growth, and industry concentration," working paper, 2021.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and misallocation in general equilibrium," *Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

Barkai, Simcha, "Declining labor and capital shares," *Journal of Finance*, 2020, *75* (5), 2421–2463.

Bergeaud, Antonin, Gilbert Cette, and Rémy Lecat, "Productivity trends in advanced countries between 1890 and 2012," *Review of Income and Wealth*, 2016, *62* (3), 420–444.

Bessen, James, "Industry concentration and information technology," *Journal of Law and Economics*, 2020, *63* (3), 531–555.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are ideas getting harder to find?," *American Economic Review*, 2020, *110* (4), 1104–1144.

Boppart, Timo and Huiyu Li, "Productivity slowdown: Reducing the measure of our ignorance," in "The Economics of Creative Destruction," Harvard University Press, Cambridge, MA, 2022.

Chatterjee, Satyajit and Burcu Eyigungor, "The firm size-leverage relationship and its implications for entry and business concentration," *Review of Economic Dynamics*, 2022.

Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon, "From good to bad concentration? U.S. industries over the past 30 years," *NBER Macroeconomics Annual*, 2020, *34* (1), 1–46.

Crouzet, Nicolas and Janice C Eberly, "Understanding weak capital investment: The role of market concentration and intangibles," *Proceedings of the 2018 Jackson Hole symposium*, May 2019, pp. 87–148.

David, Joel M and Venky Venkateswaran, "The sources of capital misallocation," *American Economic Review*, 2019, *109* (7), 2531–2567.

Eggertsson, Gauti B, Jacob A Robbins, and Ella Getz Wold, "Kaldor and Piketty's facts: The rise of monopoly power in the United States," *Journal of Monetary Economics*, 2021, *124*, S19–S38.

Farhi, Emmanuel and François Gourio, "Accounting for macro-finance trends: Market power, intangibles, and risk premia," *Brookings Papers on Economic Activity*, 2018, *2*, 147–250.

Fernald, John G, Robert E Hall, James H Stock, and Mark W Watson, "The disappointing recovery of output after 2009," *Brookings Papers on Economic Activity*, 2017, *1*, 1–58.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow, "How destructive is innovation?," *Econometrica*, 2019, *87* (5), 1507–1541.

Hall, Robert E, "New evidence on the markup of prices over marginal costs and the role of mega-firms in the US economy," NBER Working Paper 24574, 2018.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania, "From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share," *Econometrica*, 2022, *90* (4), 1879–1914.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg, "The Industrial revolution in services," *Journal of Political Economy Macroeconomics*, 2022.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Jiang, Xian, "Information and communication technology and firm geographic expansion," working paper, 2021.

Karabarbounis, Loukas and Brent Neiman, "The global decline of the labor share," *Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Kaymak, Baris and Immo Schott, "Corporate tax cuts and the decline of the labor share," working paper, 2020.

Kehrig, Matthias and Nicolas Vincent, "The micro-level anatomy of the labor share decline," *Quarterly Journal of Economics*, 2021, *136* (2), 1031–1087.

Klenow, Peter J and Huiyu Li, "Innovative growth accounting," *NBER Macroeconomics Annual*, 2021, *35* (1), 245–295.

Klette, Tor Jakob and Samuel S Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

Koh, Dongya, Raül Santaeulàlia-Llopis, and Yu Zheng, "Labor share decline and intellectual property products capital," *Econometrica*, 2020, *88* (6), 2609–2628.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard, "Information technology and returns to scale," working paper, 2021.

Liu, Ernest, Atif Mian, and Amir Sufi, "Low interest rates, market power, and productivity growth," *Econometrica*, 2022, *90* (1), 193–221.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The rise of market power and the macroeconomic implications," *Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

Luttmer, Erzo GJ, "On the mechanics of firm growth," *Review of Economic Studies*, 2011, *78* (3), 1042–1068.

Peters, Michael and Conor Walsh, "Population growth and firm-product dynamics," NBER Working Paper 29424, 2022.

Ridder, Maarten De, "Market power and innovation in the intangible economy," working paper, 2021.

Syverson, Chad, "Challenges to mismeasurement explanations for the US productivity slowdown," *Journal of Economic Perspectives*, 2017, *31* (2), 165–186.