A Theory of Falling Growth and Rising Rents

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

March 2019

Abstract

Growth has fallen in the U.S., while firm concentration and profits have risen. Meanwhile, labor's share of national income is down, mostly due to the rising market share of low labor share firms. We propose a theory for these trends in which the driving force is falling firm-level costs of spanning multiple markets, perhaps due to accelerating ICT advances. In response, the most efficient firms spread into new markets, thereby generating a temporary burst of growth. Because their efficiency is difficult to imitate, less efficient firms find their markets more difficult to enter profitably and innovate less. Even the most efficient firms do less innovation eventually because they are more likely to compete with each other if they try to expand further.

*Aghion: Collège de France and London School of Economics; Bergeaud: Banque de France; Boppart: IIES, Stockholm University; 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, the Bank of France or the Eurosystem.

1. Introduction

Recent studies have documented the following patterns in the U.S. economy over the past several decades:[1]

1. Falling "long run" growth (interrupted by a temporary burst of growth)
2. Falling labor share due to rising revenue shares of low labor share firms
3. Rising firm concentration within industries

In this paper we construct a theory of endogenous growth with heterogeneous firms which speaks to these facts. There are two main sources of heterogeneity in our model. First, *product quality* which improves endogenously on each product line through innovation and creative destruction. Second, *process efficiency*, which is time-invariant and is unevenly distributed across firms in the economy. High process efficiency firms command a higher markup than low productivity firms, conditional on having the same product quality advantage over their competitors.

A possible source of persistent differences in process efficiency across firms is their organizational capital. Firms such as Walmart and Amazon have established successful business models and logistics that are evidently hard to copy. Both firms experienced considerable expansion into new geographic and product markets over the past two decades. 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 is more observable.

Our story is that the ICT (Information and Communication Technology) wave in the 1990s has allowed high productivity firms to extend their boundaries — to expand over a wider set of product lines. We model the ICT

[1] We discuss papers presenting evidence on these patterns in the next section.

wave as a downward shift in the overhead cost $C(n)$ of running n product lines. This cost is assumed to be convex in n, which puts a brake on the quality innovation (creative destruction) efforts of high process efficiency firms. Since high productivity firms enjoy higher profits per product line, while sharing the same overhead cost function and cost of R&D inputs as low productivity firms, the downward shift in the overhead cost schedule will allow high productivity firms to expand to a higher fraction of lines. The expansion of high productivity firms fuels a temporary surge in aggregate productivity growth — both because they innovate to take over more markets (bringing quality improvements) and because they apply their superior process efficiency to those additional markets.

Since high productivity firms have higher markups and lower labor shares on average across their product lines, their expansion into more lines will result in an increase in the aggregate markup and a reduction of the aggregate labor share. This is entirely driven by firm composition rather than within-firm changes. Within-firm average markups can actually fall, as the quality leader on a product line is more likely to face a high process efficiency incumbent, limiting their markups whether they are a high or low process efficiency firm.

The expansion of high productivity firms into more lines eventually deters innovation. This is because innovating on a line where the incumbent firm has high productivity yields lower profits than when the incumbent firm is a low productivity firm. This results in lower within-firm markups. Both high and low productivity firms eventually curtail their efforts at creative destruction, knowing they will face stiffer competition. This in turn lowers long run innovation and productivity growth. The drop in long run growth leads to a lower pace of job reallocation, which is tied to creative destruction.

We calibrate our model to gauge the strength of the model's mechanism. We choose parameter values to fit the % point decline in labor share (and its firm composition component), the rise in firm concentration, "initial" productivity growth (before the drop in overhead costs), and an initial real interest rate, and rate of investment in R&D. The model can plausibly generate around one half of the slowdown growth seen in recent years.

Most closely related to our paper are Akcigit and Ates (2018) and Liu, Mian and Sufi (2019), who both study declining growth and rising concentration. We differ from these papers by emphasizing the ICT wave and its effect on high productivity versus low productivity firms as a main driving force. They emphasize, instead, changes in competition policy or the decline in the cost of credit as main driving forces. Unlike these two papers, we generate opposite trends for labor share (and markups) within versus across firms. Finally, our model generates a burst of growth before a long run growth slowdown.

Our paper also relates to Chatterjee and Eyigungor (2017) which studies the rise in concentration; to recent papers on declining labor share, in particular Karabarbounis and Neiman (2013, 2018), Barkai (2016), Eggertsson, Mehrotra, Singh and Summers (2016), Autor, Dorn, Katz, Patterson and Van Reenen (2017), Martinez (2017), and Farhi and Gourio (2018). Whereas Autor et al. (2017) looked at labor share in U.S. Census data, Baqaee and Farhi (2017) and De Loecker and Eeckhout (2017) estimate markups in Compustat firms. These latter papers decompose the recent evolution of of the aggregate markup into within-firm and between-firm components. They find the dominant contributor to be the rising market share of high markup firms. We contribute to this literature by providing a theoretical framework that links the rise in concentration and the rise in average markups (similarly, decline in aggregate labor share) to the slowdown in U.S. growth in recent decades.

The rest of the paper is organized as follows. Section 2. describes the empirical patterns documented by other studies that motivate our modeling effort. Section 3. lays out our model, and solves analytically for its steady state. Here we qualitatively characterize the steady state comparative statics due to falling overhead costs. In section 4. we calibrate the model to see whether it can generate a realistic decline in long run growth. Section 5. concludes.

2. Stylized facts

Fact 1: Falling "long run" growth (after a burst of growth) Figure 1 presents U.S. annual TFP growth over subperiods from the U.S. Bureau of Labor Statistics (BLS). Note that 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 intellectually property investments; we consistently included this portion in residual TFP growth as part of what we are trying to explain.

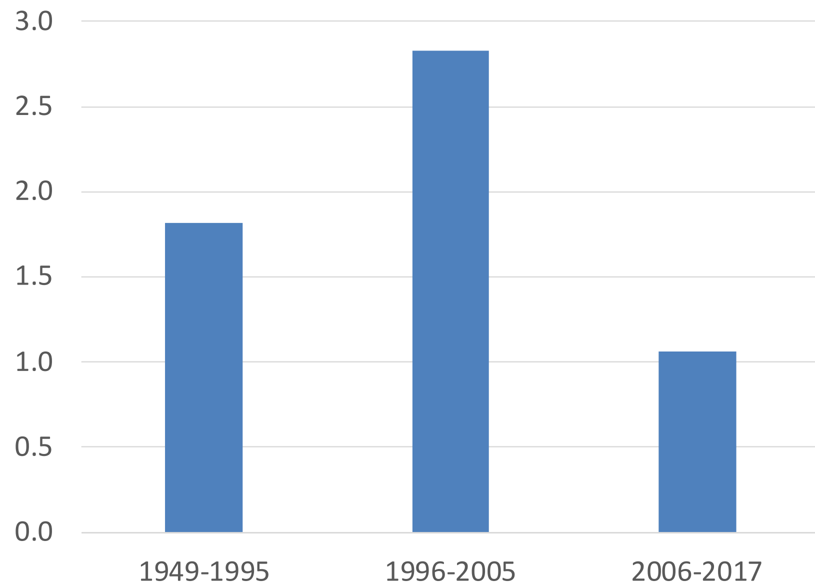
The Figure shows growth accelerating from its 1949–1995 average of 1.8% per year to 2.8% per year from 1996–2005, before falling to just 1.1% per year from 2006–2017. Fernald, Hall, Stock and Watson (2017) argue that the recent slowdown is statistically significant and predates the Great Recession. Syverson (2017) and Aghion, Bergeaud, Boppart, Klenow and Li (2019) contend that the slowdown is real and unlikely to be fully attributable to growing measurement errors.

Fact 2: Falling labor share (mostly due to composition) Figure 2 shows that, according to the BLS, the aggregate U.S. labor share of output in the nonfarm business sector fell about six % points in the last two decades.[2] Autor, Dorn, Katz, Patterson and Van Reenen (2017) find declining labor share in a number of Census sectors, but most sharply in manufacturing. Table 1 reproduces their statistics on the cumulative change in labor share for six Census sectors in recent decades. Finance is the contrarian, with rising labor share. In five of the six sectors the sales shares shifted to low labor share firms, so that the "between" component pushed labor share downward notably. And within-firm labor shares actually rose in all sectors but manufacturing.

In the business cycle literature, labor share is often used as an inverse measure of price-cost markups. See Karabarbounis (2014) and Bils, Klenow

[2] Since this is the business sector, it is not affected by the Rognlie (2016) critique that the rise of housing is exaggerating the decline in labor share.

Figure 1: U.S. productivity growth rate

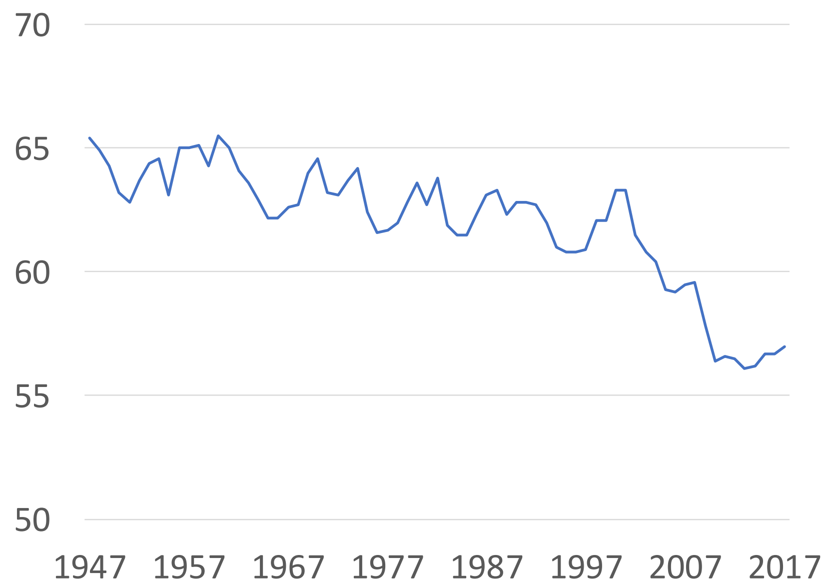


Source: BLS multifactor productivity series. We calculate yearly productivity growth rate by adding R&D and IP contribution to BLS MFP and then converting the sum to labor augmenting form. The figure plots the average productivity growth within each subperiod. The unit is percentage points.

and Malin (2018). Thus one interpretation of falling labor share due to composition effects is that markups are rising due to composition effects. De Loecker and Eeckhout (2017) and Baqaee and Farhi (2017) do just that, though with a broader measure of variable inputs that adds intermediates to labor costs. A competing interpretation is that the elasticity of output with respect to capital has risen. Barkai (2016), Gutiérrez and Philippon (2016, 2017), and Farhi and Gourio (2018) argue against this interpretation and in favor of rising markups on the grounds that the investment rate and capital-output ratio have not risen.[3]

[3]Koh, Santaeulalia-Llopis and Zheng (2016) and Traina (2018) argue that labor share has not fallen and markups have not increased if one adds intangibles investment such R&D and marketing. But these expenditures are arguably not part of variable costs.

Figure 2: U.S. labor share



Source: BLS. The figure plots the aggregate labor compensation of all employed persons as a share of aggregate output for the nonfarm business sector. The unit is percentage points.

Table 1: Cumulative change in labor share over given period (ppt)

	1982–2012				92–12	92–07
	MFG	RET	WHO	SRV	FIN	UTL
$\Delta\frac{\text{Payroll}}{\text{Sales}}$	-7.01	-0.79	0.19	-0.19	3.25	-1.89
Within	-1.19	3.74	4.01	2.43	6.29	0.58
Between	-4.97	-4.03	-4.38	-0.44	-3.62	-2.39

Source: Table 5 in Autor et al. (2017). This is a Melitz-Polanec decomposition of the change in the labor share. The unit is percentage points.

Fact 3: Rising concentration Table 2, which is also based on Autor, Dorn, Katz, Patterson and Van Reenen (2017), presents the average 5-year change in top 4 or top 20 firm concentration ratios in 4-digit NAICS. These results are, again, from firm-level data in U.S. Census years. Across the six sectors, the top 4 firm shares increase from 0.4 to 2.5 percentage points per five-year period, while the top 20 firm shares increase between 0.8 and 3.6 percentage points per year. Concentration increased the most rapidly in retail and finance, and least rapidly in manufacturing and wholesale.

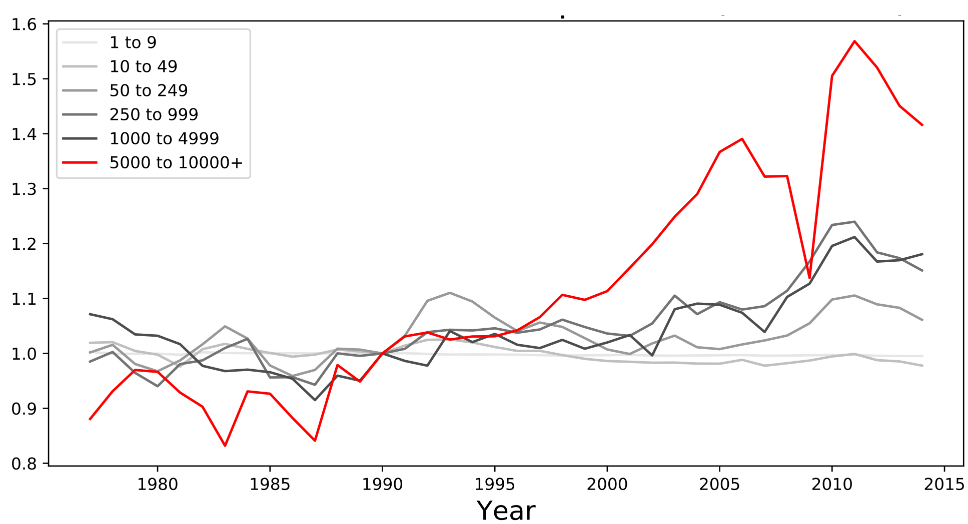
The rise in concentration in Table 2 is at the national level. In contrast, Rossi-Hansberg, Sarte and Trachter (2018) and Rinz (2018) find that local concentration declined. One explanation for the diverging trends is that the largest firms grew by adding establishments in new locations. Figure 3 shows cumulative growth of the number of establishments per firm, by firm size bins, in the Business Dynamic Statistics from the Census Bureau. The red line is the growth of establishments for the largest firms. It shows that, between 1990 and 2014, the largest firms expanded by adding establishments. The average number of establishments rose for smaller firms too but not as quickly as for the largest firms. In a parallel study, Cao, Mukoyama and Sager (2018) document a similar pattern in the Quarterly Census of Employment and Wages data. To the extent that growth in the number of establishments is connected to growth in the number of products or markets, this evidence suggests that the rise in national concentration may not reflect an increase in market power of the largest firms.

Table 2: Average 5-year change in national concentration (ppt)

	1982–2012				92–12	92–07
	MFG	RET	WHO	SRV	FIN	UTL
Δ Top 4 firms sales share	0.7	2.5	0.4	0.7	2.1	1.9
Δ Top 20 firms sales share	0.8	2.7	1.0	1.0	3.6	1.2

Source: Table 1 of Autor et al. (2017). Averages across 4-digit industries, with the industries weighted by industry sales shares.

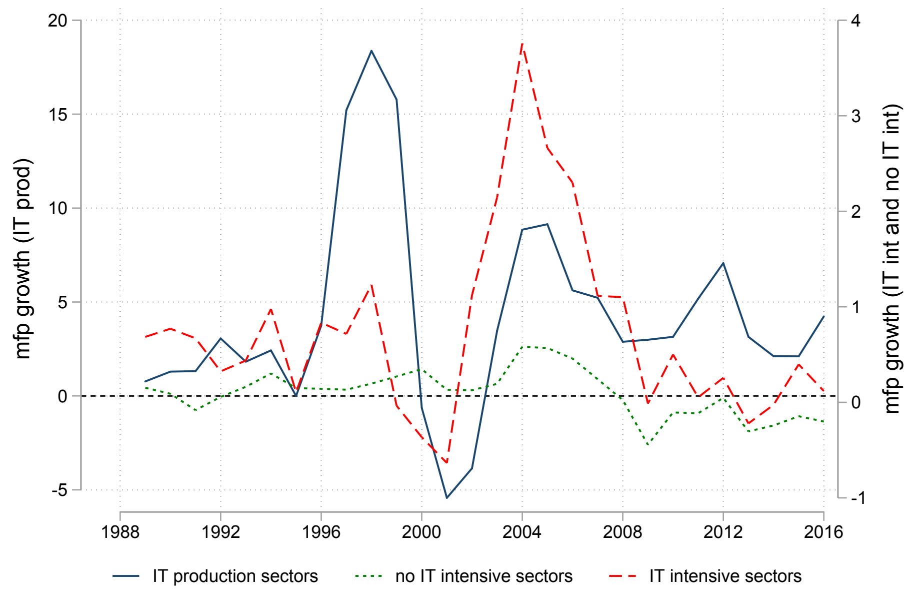
Figure 3: Growth in establishment per firm by firm size



Source: U.S. Census Bureau Business Dynamic Statistics. The graph plots the number of establishments per firm within employment bins relative to 1990

ICT as a driving force We focus on changes in ICT as a possible driver of the patterns described above for three reasons. First, Figure 4 displays growth rate of multi-factor productivity for IT-producing, IT-intensive and non-IT-intensive industries, classified based on Fernald (2015). The figure shows a

Figure 4: Productivity growth by ICT intensity rate



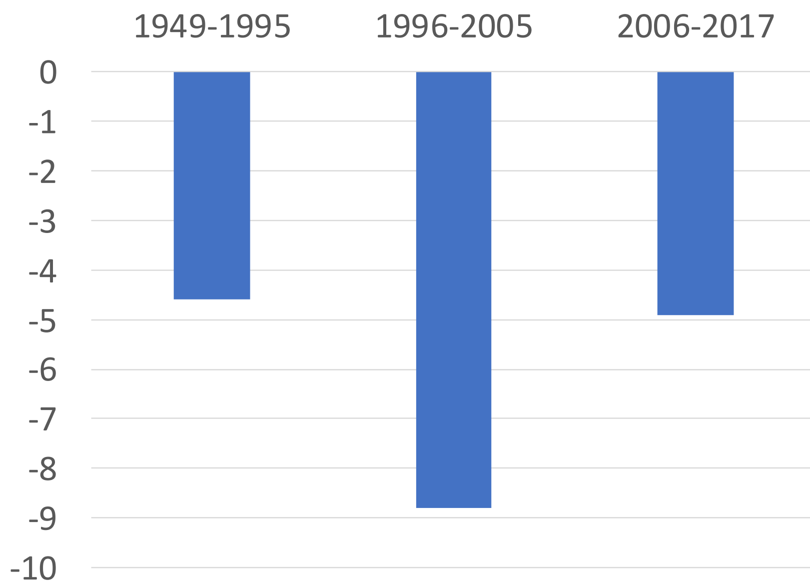
Source: updated version of Fernald (2014) Fig 6A. % per year, 4 year MA.

burst of growth for the IT-intensive sectors in early 2000's after a burst of growth for IT-producing sectors in the second half of 1990. In contrast, the non-IT-intensive sectors did not experience a burst of growth.

Second, price declines for ICT goods accelerated sharply for a decade from the mid-1990s to the mid-2000s. See Figure 5. This is in the middle of the period of rising concentration.

Third, Crouzet and Eberly (2018) and Bauer, Boussard and Laskhari (2018) document that bigger firms invest a higher share of their sales in intangibles and ICT, respectively. Lower costs of ICT would seem to benefit larger firms more. The former evidence is for U.S. firms and the latter for French firms.

Figure 5: Relative price of ICT



Source: BEA. % Change per year in the price of ICT relative to the GDP deflator.

3. Model

As motivated by the empirical micro evidence behind the trend in aggregate labor income share the data asks for a model of firm heterogeneity. Furthermore in order to make a prediction regarding the more recently observed productivity slowdown a model of endogenous growth is needed. Our strategy is to lay out a simple such framework that is however rich enough to speak to the facts in Section 2. The goal is to build a parsimonious theoretical building block that can straightforwardly be augmented in various ways and then still serve as a backbone in much richer frameworks. We discuss various extensions and generalizations of our framework in order to make the theory more quantitative. However, we also show in a first calibration that even tractable baseline model has already some quantitative bite. This simple calibration does generate significant quantitative effects in line with the trends observed in the U.S. over the past 30 years.

3.1. Preferences

The household side is relatively standard. Time is discrete and the economy is populated by a representative household who chooses consumption C to maximizes the preferences

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

subject to

$$a_{t+1} = (1 + r_t)a_t + w_t L - C_t, \tag{1}$$

a standard no-Ponzi game condition and a given initial wealth level $a_0 > 0$. Here a denotes wealth, r the interest rate, w the wage rate and L is the labor endowment that is inelastically supplied to the labor market.

The Euler equation resulting from household's optimization is given by

$$\frac{C_{t+1}}{C_t} = \beta(1 + r_{t+1}). \tag{2}$$

3.2. Production of final output

A final output good is produced competitively out of a unit interval of intermediate inputs according to the following Cobb-Douglas technology

$$Y = \exp\left(\int_0^1 \log\left[q(i)y(i)\right]di\right). \tag{3}$$

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

$$y(i) = \frac{YP}{p(i)}, \tag{4}$$

where we have the aggregate price index

$$P = \exp\left(\int_0^1 \log\left[p(i)/q(i)\right] di\right), \tag{5}$$

which we will in the following normalize to one in each period.

3.3. Production and market structure for intermediate inputs

There is an interval of firms indexed by $j \in [0, J]$. 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) heterogeneity in the firm-market specific quality $q(i, j)$ which evolves endogenously as a result of innovation in the quality dimension; (ii) heterogeneity in a permanent firm-specific level of process efficiency.

We first describe the heterogeneity in the process efficiency. There is a firm specific level of process efficiency denoted by $\varphi(j)$. A firm with process efficiency $\varphi(j)$ can produce in any line i with the following simple linear technology

$$y(i, j) = \varphi(j)l(i, j), \tag{6}$$

where $l(i, j)$ denotes labor used by firm j to produce in line i and $y(i, j)$ denotes the output of this firm in this line. We assume that the heterogeneity in process efficiency is persistent over time reflecting, e.g., differences in organizational capital that is hard to replicate. This heterogeneity in process efficiency will lead in our model to persistent differences in revenue TFP and labor income shares across firms.

Note that the linear technology in (6) applies irrespective of the specific quality firm j produces in line i. In addition to the heterogeneity in process efficiency, firms differ in the quality at which they can produce in a line i. We will below explain how the distribution of quality of the different lines across the firms changes endogenously due to innovations. But for the static firm problem here we just assume that in a period t each firm j can produce at some line-specific quality $q(i,j)$. Labor is fully mobile such that the wage rate equalizes across firms. Hence, the marginal cost of firm j per output unit of line i is given by $w/\varphi(j)$, or the marginal cost per quality-adjusted unit of line i, $q(i,j)y(i,j)$ is equal to $\frac{w}{q(i,j)\varphi(j)}$.

3.4. Pricing

In each market all the firms engage in Bertrand competition. This means that only the firm with the highest quality-adjusted productivity $q(i,j)\cdot\varphi(j)$ is active in a given market i and sets the price such that the firm with the second highest quality-adjusted productivity finds it not profitable to be active. In the following we denote the index of the leading firm in line i by $j(i)$ and the second-best firm by $j'(i)$. Hence, the quality-adjusted productivity of the "leader" is given by $q(i,j(i))\cdot\varphi(j(i))$ and the same quality-adjusted productivity of the second-best firm in line i is given by $q(i,j'(i))\cdot\varphi(j'(i))$. The price setting behavior of this leader is constrained by the second-best firm and the leader will set its quality-adjusted price equal to the quality-adjusted marginal cost of the second-best firm. Formally we then have

$$\frac{p(i,j(i),j'(i))}{q(i,j(i))} = \frac{w}{q(i,j'(i))\cdot\varphi(j'(i))}. \tag{7}$$

Note that the price charged in line i depends on the process efficiency of both the leader and the follower as well as the quality difference between the two. We define the markup in line i, $\mu(i)$, as the ratio between the price of a unit divided by the marginal cost of the producer. The markup is then given by

$$\mu(i,j(i),j'(i)) \equiv \frac{p(i,j(i),j'(i))}{w/\varphi(j(i))} = \frac{q(i,j(i))\cdot\varphi(j(i))}{q(i,j'(i))\cdot\varphi(j'(i))}. \tag{8}$$

The markup of a product increases in the quality gap $\frac{q(i,j(i))}{q(i,j'(i))}$ as well as in the process efficiency gap $\frac{\varphi(j(i))}{\varphi(j'(i))}$ between the leader and the second-best firm. In particular, all else equal, product level markup is increasing with the process efficiency of the leading firm $\varphi(j(i))$ and decreasing in the process efficiency of the second-best firm $\varphi(j'(i))$. Within a firm the markup differs across product lines but a firm with a higher process efficiency will charge on average a higher markup.

Given the markup $\mu(i)$ the operating profits the leader makes in line i follows directly. Combining the pricing with the demand (4), using the definition of the numéraire gives for the operating profits in a period

$$\pi(i) = Y\left(1 - \frac{1}{\mu(i)}\right). \tag{9}$$

3.5. Innovation and productivity growth

The quality distribution evolves endogenously as a result of innovation. Any firm j can engage in R&D activity to acquire a patent to produce in a new line at higher quality . More specifically, by investing $x_t(j)\psi_c Y_t$ units of final output in period t in R&D $x_t(j)$ product lines are randomly drawn from 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.

The initial distribution of quality levels in the different lines across firms is exogenously given. In each line the firm with the highest quality will face competition by a firm with lower quality by a factor γ.

We assume that a period is short enough such that no two innovations arrive in the same line in a given period. If we denote the innovation rate of firm j in period t by $x_t(j)$ the aggregate rate of creative destruction is given by

$$z_{t+1} = \int_0^J x_t(j)dj, \tag{10}$$

i.e., for any given line an innovation arrives in period $t+1$ with probability z_{t+1}.

This endogenous quality improvement due to creative destruction is the origin of long-run growth in this model.

3.6. Markups with binary process efficiency levels

For simplicity we assume in the following two types of firms. Firms indexed by $j \in [0, \phi J]$ are of high process efficiency type φ_H whereas the remaining fraction $1 - \phi$ of all firms is of low process efficiency type φ_L. We denote the productivity differential by $\Delta \equiv \varphi_H / \varphi_L > 1$. So in total there are ϕJ high productivity producer and $(1 - \phi)J$ low productivity producers. In this baseline model we further assume $\gamma > \Delta$ such that the firm with the highest quality is always the active leader irrespective of whether she is of high or low process efficiency.

Then, with the two process efficiency levels (high or low) there are four potential cases of markups and profits:

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. In this case we have

$$\mu(i) = \gamma, \tag{11}$$

and

$$\pi(i) = Y\left(1 - \frac{1}{\gamma}\right) \equiv \pi_{HH}. \tag{12}$$

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. In this case we have

$$\mu(i) = \Delta\gamma, \tag{13}$$

and

$$\pi(i) = Y\left(1 - \frac{1}{\Delta\gamma}\right) \equiv \pi_{HL}. \tag{14}$$

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. In this case we have

$$\mu(i) = \frac{\gamma}{\Delta}, \tag{15}$$

and

$$\pi(i) = Y\left(1 - \frac{\Delta}{\gamma}\right) \equiv \pi_{LH}. \tag{16}$$

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. In this case we have

$$\mu(i) = \gamma, \tag{17}$$

and

$$\pi(i) = Y\left(1 - \frac{1}{\gamma}\right) \equiv \pi_{LL}. \tag{18}$$

3.7. Boundary of the firm

Given the constant cost of acquiring a line through innovation and the fact that firms with a high process efficiency make higher expected operating profits in an additional line, high productivity firm have ceteris paribus a higher incentive to undertake R&D activity. To prevent the high productivity firms from taking over all lines and to have a well defined boundary of the firm, we assume that firms have to pay an additional per-period overhead cost which is a convex function of the number of lines operated. More specifically, suppose the number of lines firm j operates is denoted by $n(j)$. We assume a quadratic per-period overhead cost in terms of final output in this number of lines $n(j)$, namely

$$\frac{1}{2}\psi_o n(j)^2 Y, \tag{19}$$

with $\psi_o > 0$. The convexity of the overhead cost in the number of product lines $n(j)$, gives rise to a natural boundary of the firm. High productivity firms will typically operate more lines than low productivity firms, but no firm (type) will operate all lines.

It may be worthwhile to briefly compare our model to Klette and Kortum (2004) which serves as a benchmark in this literature. Here we assume a linear cost of innovating on a new line and convex overhead cost. Consequently, the (expected) marginal value of an additional line in a firm is decreasing in the number of lines, $n(j)$, and this diminishing marginal value defines a natural boundary of the firm. By contrast, Klette and Kortum (2004) assumes a convex cost of acquiring extra product lines through creative destruction, and a non-diminishing value of additional lines. Our modelling approach allows us to do comparative statics with respect to the scalar ψ_o which affecting the boundary of firms without altering the technology to undertake innovations. In our main application we argue that ICT improvements lower ψ_o permanently (for all firms) and we then study its effect on concentration, labor share and growth during the transition as well as in the new steady state. Another difference with Klette and Kortum (2004), is that we assume here that firms operate on a continuum of lines, so that the law of large numbers applies. One consequence is that there is no firm exit in our baseline model.[4]

3.8. Labor income shares

This simple version of the model abstracts from physical capital and labor is the only factor in variable production. Furthermore we assumed both R&D expenditure and overhead costs are denominated in final output. These last two assumptions are made to avoid a mechanical effect of the firm size distribution (and overhead cost) and the overall level of R&D activity on the labor income share. Hence in this framework the aggregate labor income share is simply determined by the distribution of markups across lines.

Because of the Cobb-Douglas technology in final output production, the revenue from each product is equal to Y. Then the total variable cost in a line i

[4] In a model extension we consider (gross) entry and exit too.

is equal to

$$wl(i) = \frac{Y}{\mu(i)}.$$

Integrating both sides of the above equation over all i yields

$$wL = Y \int_0^1 \frac{1}{\mu(i)} di.$$

Dividing the above two equations by each other we then get the cost (or employment) share of product line i as

$$\frac{l(i)}{L} = \frac{1}{\mu(i)} \frac{1}{\int_0^1 \frac{1}{\mu(\iota)} d\iota}. \tag{20}$$

The relative cost per line, $l(i)/L$, is inversely proportional to the markup factor per line. This comes from revenue being equalized across lines due to the Cobb-Douglas technology in final production.

Finally, the aggregate labor income share is given by

$$1 - \alpha \equiv \frac{wL}{Y} = \frac{1}{\int_0^1 \mu(i) l(i)/L \, di} = \int_0^1 \mu(i)^{-1} \, di. \tag{21}$$

This is identical to the inverse of the average cost-weighted markup factor. There is no physical capital in this model and the profit share and the labor income share add up to one. However, the aggregate labor share depends non-trivially upon the full distribution of markups across lines. This distribution is determined by the type of the leader and second-best firm across lines.

What about the labor income share at the firm level? Consider firm j with $n(j)$ lines that faces in a fraction $h(j)$ of these lines a high type second-best firm and in the remaining fraction $1-h(j)$ a low productivity second-best firm. Note that the firm's labor income share in a line i is simply given by $\frac{1}{\mu(i)}$. If firm j is itself of high type, its overall labor income share is given by

$$1-\alpha(j)=h(j)\frac{1}{\gamma}+(1-h(j))\frac{1}{\gamma\Delta}. \tag{22}$$

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

$$1-\alpha(j)=h(j)\frac{\Delta}{\gamma}+(1-h(j))\frac{1}{\gamma}. \tag{23}$$

Faced by the same share of high type competitors $h(j)$, firms with a higher process efficiency have a lower labor income share (as they can on average charge a higher markup). Hence the model will generate persistent differences in the labor income share across firms[5]. However, since the composition of competitors $h(j)$ is endogenous the model is flexible enough to also generate changes over time in the labor income share *within* firms.

3.9. Dynamic firm problem

There are two individual state variables in the firm problem: the number of lines a firm j operates, $n(j)$, and the fraction of high productivity second-best firms firm j faces in these lines $h(j)$. Each firm then chooses in how many new lines to innovate x_t to maximize the net present value of future profits. Let us denote by Π^H and Π^L the period profits of a firm after overhead relative to total output, Y, of a high and low type firm operating $n(j)$ and facing competition of high second-best firms in a fraction of $h(j)$ of them. Formally, we have

$$\Pi^H(n(j),h(j))=n(j)-\frac{n(j)h(j)}{\gamma}-\frac{n(j)\,(1-h(j))}{\gamma\Delta}-\frac{1}{2}\psi_o n(j)^2,$$

and

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

[5] See Hsieh and Klenow (2009) and David and Venkateswaran (forthcoming) for evidence on persistent difference in revenue per worker.

Note since these are profits divided by output Y_t these profit shares in GDP are time invariant. The problem of a firm of type $j = H, L$ can be written as

$$V_0 = \max_{\{x_t, n_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} Y_t \left[\Pi^j(n_t, h_t) - x_t \psi_c\right] \prod_{s=0}^{t} \left(\frac{1}{1+r_s}\right) \tag{24}$$

subject to

$$n_{t+1} = n_t(1 - z_{t+1}) + x_t, \tag{25}$$

$$h_{t+1} n_{t+1} = h_t n_t (1 - z_{t+1}) + S_t x_t, \tag{26}$$

and a given initial n_0 and h_0. The constraint (25) states that the number of product lines of a firm tomorrow is equal to the newly added lines x plus the number of lines today times one minus the rate of creative destruction in the economy, z. The second constraint (26) states that the number of lines in which the firm faces a high type second-best firm is equal to the number of such lines today times $1 - z$ plus the number of newly added lines times the average fraction of lines operated by high productivity type firms S_t. When optimizing the firm takes the path of output Y_t, the interest rate r_t, the rate of creative destruction z_{t+1}, and the aggregate fraction of lines operated by high productivity firms S_t as given.

3.10. Market clearing and resource constraints

We close the model with the following market clearing conditions. First, final output will be used for consumption, C, total overhead cost, O, and total R&D expenditure, Z, or formally

$$Y = C + O + Z, \tag{27}$$

where

$$O = \int_0^J \frac{1}{2} \psi_o n(j)^2 Y dj, \tag{28}$$

and

$$Z = \int_0^J x(j)\psi_c Y dj. \tag{29}$$

Labor is only used as a variable input by the producer of different intermediate inputs. Labor market clearing implies

$$L = \int_0^J \int_0^1 l_t(j,i)\, di\, dj,$$

where $l_t(j,i)$ denotes labor used by firm j that operates i. Furthermore, asset market clearing requires

$$\int_0^J V_t(j)dj = a_t. \tag{30}$$

In addition, we have the equations defining the aggregate share of lines operated by high types as

$$S_t = \int_0^{\phi J} n_t(j)dj, \tag{31}$$

an accounting equation that states that all lines are operated by a firm

$$1 = \int_0^J n_t(j)dj, \tag{32}$$

and a final equation that relates output to the distribution of process efficiency, quality levels and markups

$$Y_t = Q_t \frac{\varphi_L \Delta^{S_t} \exp\left[-\int_0^1 \log\left(\mu_t(i)\right) di\right]}{\int_0^1 \left(\mu_t(i)\right)^{-1} di} L, \tag{33}$$

where $Q_t = \exp\left[\int_0^1 \log\left(q_t(i,j)\right) di\right]$ denotes the "average" quality level. An equilibrium in this economy is then a path of allocation and prices that jointly solves the household problem, the firms' problems and is consistent with the market clearing and accounting equations stated above.

Since the output level is a function of the full distribution of markups across product lines the equilibrium path is in general a function of the entire initial distribution of product lines $n(j)$ and competition $h(j)$ across all firms. One approach to simplify this is to assume that all firms of the same type start our with the same level of $n(j)$ and $h(j)$. Since the law of large number applies the firms of the same type will be identical along the entire equilibrium path and just two firm problems—one for a high type and one for a low type—need to be solved. The aggregate state vector can then be summarized by the three variables S, and the share of high second-best firms in lines operated by the two types h_L and h_H.

In the following we however do not analyze the transitional path but rather focus on the steady state. We will show below that this steady state can be fully characterized in this economy and takes very tractable functional forms. We then discuss how a permanent drop in ψ_o (triggered by improvements in ICT) affects market concentration, labor income shares (within firms as well as on the aggregate), and productivity growth in the long run.

3.11. Steady state definition

In the following we focus on the stationary equilibrium this economy will asymptotically converge to. We define this steady state equilibrium in the following way:

Definition 1 *A steady state is an equilibrium path along which the interest rate is constant and equal to $r^\star$ and a constant fraction of the lines , $S^\star$, is provided by high productivity producers.*

In a steady state the number of products must be equalized across all firms of the same type and must be constant over time. The number of products however will differ between firms of different types. So all high productivity type firms have $n(j)^\star = n_H^\star$ and all low type firms have $n(j)^\star = n_L^\star$. For a constant number of lines within firm, the R&D activity of each firm must be proportional to its number of products, i.e., $x(j)^\star = n(j)^\star z^\star$. Since all firms draw new lines from a stationary distribution, the fraction of high productivity type second-best firms faced is equalized across firms and we have

$$h(j)^{\star} = S^{\star}, \forall j. \tag{34}$$

Since the markup distribution is stationary in steady state output Y_t is proportional to the average quality Q_t (see equation (33)). Consequently we have

$$\frac{Y_{t+1}}{Y_t} = \frac{Q_{t+1}}{Q_t} = \gamma^{z^{\star}} \equiv g^{\star}. \tag{35}$$

Finally since total overhead, O, total R&D expenditure, Z, all grow at the same gross rate $g^{\star}$ also consumption has to grow at this rate $g^{\star}$ (see (27)). Then, the Euler equation determines the steady state interest rate as

$$r^{\star} = \frac{g^{\star}}{\beta} - 1. \tag{36}$$

Next, we formally characterize this steady state and analyze its solution.

3.12. Steady state characterization

Let us denote by v the value of a firm V relative to total output Y. In steady state (with $h(j)^{\star} = S^{\star}$) the number of products per firm becomes the only state variable and we can write $v(n)$. All high productivity then solve

$$v_H(n) = \max_{n'} \{\Pi^H(n, S^{\star}) - (n' - n(1 - z^{\star}))\psi_c + \beta v_H(n')\}.$$

Similarly, all low productivity firms solve

$$v_L(n) = \max_{n'} \{\Pi^L(n, S^{\star}) - (n' - n(1 - z^{\star}))\psi_c + \beta v_L(n')\}.$$

The two accounting equations (30) and (31) become in steady state

$$S^{\star} = n_H^{\star}\phi J \tag{37}$$

and

$$n_H^\star \phi J + n_L^\star (1-\phi) J = 1. \tag{38}$$

Finally in steady state we must have

$$n_H^\star = f_H(n_H^\star) \tag{39}$$

and

$$n_L^\star = f_L(n_L^\star) \tag{40}$$

where $f_H(\cdot)$ and $f_L(\cdot)$ are the policy functions of the high and low types. These equations fully characterize the steady state.

Since the π^H and π^L are quadratic functions of n this is a very simple dynamic programming problem for which the policy functions become linear and the value function can be characterized in closed form. We state the solution of the dynamic programming problems in the next proposition.

Proposition 1 *The steady state is a $(n_H^\star, n_L^\star, S^\star, z^\star)$ combination that fulfills*

$$\phi J n_H^\star = S^\star \textit{ and } (1-\phi) J n_L^\star + \phi J n_H^\star = 1, \tag{41}$$

as well as the following research arbitrage equations for high and low productivity firms respectively:

$$\psi_c = \frac{1 - S^\star/\gamma - (1-S^\star)/(\gamma\Delta) - \psi_o n_H^\star}{1/\beta - 1 + z^\star}; \tag{42}$$

$$\psi_c = \frac{1 - S^\star\Delta/\gamma - (1-S^\star)/\gamma - \psi_o n_L^\star}{1/\beta - 1 + z^\star}. \tag{43}$$

Proof. The first-order condition of the firm problem of the high type is simply

$$\psi_c = \beta \frac{\partial v_H(n')}{\partial n'}. \tag{44}$$

Now, using the envelope theorem we get

$$\frac{\partial v_H(n)}{\partial n} = 1 - S^\star/\gamma - (1 - S^\star)/(\gamma\Delta) - \psi_o n + (1 - z^\star)\psi_c \tag{45}$$

Updating this equation, evaluating it at $n = n' = n_H^\star$ and combining it with (44) then immediately yields the research arbitrage equation of the high type. The research arbitrage equation of the low type is derived in an analog way. ■

The intuition of the two research arbitrage equations is straightforward: The optimality condition states that the marginal cost of innovating in a line, ψ should be equal to the marginal (expected) value of having an additional line. This marginal value consists in the case of the high type of the marginal profit $1 - S^\star/\gamma - (1 - S^\star)/(\gamma\Delta)$ minus the marginal overhead cost $\psi_o n$ divided by the denominator $1/\beta - 1 + z^\star$ because there is time discounting and because there is a probability $z^\star$ of loosing the additional line again in each future period.

Equations (41)–(43) are four equations in the four unknowns $(n_H^\star, n_L^\star, S^\star, z^\star)$ and can be solved explicitly. All the other endogenous variables then follow immediately. We have the following steady state results:

Proposition 2 *In steady state (i) the share of lines operated by high productivity firms is equal to*

$$S^\star = \phi + \frac{(\Delta - 1)\phi + 1}{(\Xi - 1)(\Delta - 1)}, \tag{46}$$

where $\Xi \equiv \frac{\psi_0}{J} \frac{\Delta\gamma}{(\Delta-1)^2\phi(1-\phi)}$.

(ii) High productivity firms operate more lines than low productivity types, i.e.,

$$n_H^\star > n_L^\star. \tag{47}$$

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

$$1 - \alpha_H^\star = S^\star \frac{1}{\gamma} + (1 - S^\star)\frac{1}{\gamma\Delta} \tag{48}$$

which is strictly larger than the labor income share of a low type firm is given by

$$1-\alpha_L^\star = S^\star\frac{\Delta}{\gamma} + (1-S^\star)\frac{1}{\gamma}. \tag{49}$$

Proof. The solution for $S^\star$ follows immediately from the system (41)–(43). For the difference in the number of products we get

$$n_H^\star - n_L^\star = \frac{(S^\star - \phi)}{J\phi(1-\phi)} > 0.$$

The labor income shares follow from (22), (23) and (34). ■

In this tractable model one can explicitly solve for $S^\star$. In steady state $S^\star$ can be viewed as a summary statistic of market concentration. One result that is worthwhile to note is that all the endogenous steady state variables only depend on the ratio $\frac{\psi_o}{J}$ and not on the individual level of ψ_o or J.

High process efficiency firms can (on average) charge higher markups. Consequently their incentive to undertake R&D is higher and they run into a steeper area of the convex overhead cost, i.e., operate more line than low process efficiency firms. A corollary of this is that is that high productivity firms have are larger in sales than low productivity firms ($S^\star > \phi$). They are also larger in employment although the employment difference is smaller than the sales difference because high productivity firms charge higher markups and hence have lower labor share.

3.13. Steady state effects as ψ_o decreases

In this section we do a simple comparative static exercise: Suppose improvements in ICT decrease ψ_o permanently to a lower level. How does the new steady state compare to the old one? We are particularly interested in the changes in the following endogenous variables: (i) market concentration, $S^\star$, (ii) the labor income share at the aggregate level as well as within firm, (iii) and the long-run growth rate.

The next proposition states the comparative static effect with respect to concentration.

Proposition 3 *Concentration $S^\star$ increases monotonically as ψ_o decreases.*

Proof. The comparative static effect follows directly from the expression (46).
■

The intuition that an fall in ψ_o increases $S^\star$ is the following: With a lower ψ_o a larger size gap $n_H^\star - n_L^\star$ is needed to yield the same difference in the marginal overhead cost. Consequently the high process efficiency firms will operate more lines whereas the low productivity firms shrink in size as ψ_o decreases (and market concentration goes up). This is an endogenous response in our model since the decrease in ψ_o is the same for all firms.

In the next proposition we turn to the labor income share.

Proposition 4 *As ψ_o decreases (i) the labor income share within firms increases, (ii) the reallocation of market shares goes in the opposite direction, (iii) as a consequence the aggregate labor income share may increase or decrease.*

The theory makes very sharp predictions about the labor income shares at the aggregate and micro level. As $S^\star$ decreases due to the drop in ψ_o all firms are more likely to face a high process efficiency firm as a second-best competitor in a given line. As a consequence *within* firm the labor income share *increases* (see (48) and (49)) as within firms the markup charged decreases. However, there is sales reallocation across firms that goes the opposite direction. As $S^\star$ decreases the high productive firms with a lower labor income share expand as the low productivity firms contract. This across firm effect pushes the aggregate labor income share again down. As emphasized in Section 2. these within and across firm effects that go in opposite directions are a very salient feature of the micro data.

In this simple model it is easy to show that the across firm effect dominates and the aggregate labor income share falls as ψ_o decreases if and only if $S^\star > 1/2$.

Finally, let us analyze the comparative static effects on the long-run growth rate in the next proposition.

Proposition 5 *There are also two counteracting forces on the long-run growth rate and $g^{\star}$ can either fall or increase as ψ_o drops.*

A decrease in ψ_o has a direct positive effect on the incentive to innovate: as the overhead cost decreases the marginal value of operating an additional line increases. However, there is a general equilibrium effect of the increasing $S^{\star}$ that goes in the opposite direction. With a higher $S^{\star}$ the expected markup a firm can charge in an additional line goes down and this decreases the incentive to undertake R&D expenditure. Which of the two effects dominates depends on the precise parameters. In Section 4. we show that the model indeed predicts a productivity slowdown under a first simple calibration. Our theory then predicts that this productivity slowdown should be accompanied by decreasing rate of creative destruction and consequently less churning in the labor market.

We illustrated that our tractable theory is flexible enough to speak to our motivating facts. As shown, qualitatively our theory can generate a productivity slowdown, a rising concentration, and systematic changes in the labor income shares within firms and across firms as the outcome of increasing boundary of the firm (drop in ψ_o) presumably triggered by ICT improvements. The next step is to gauge the quantitative size of these effects in a simple calibration. This we will undertake in Section 4.. In this calibration we show that even this simple theory has some quantitative bite.

However, before turning to the calibration we briefly discuss our theory's prediction along the transition (and how it will generate a burst in productivity growth along the transition) and discuss various potential generalizations and extensions of our theoretical framework.

3.14. Burst in growth during transition

So far we did a steady state comparison. In the data we saw that the productivity slowdown after the mid 2000 was preceded by a burst in productivity growth. It is indeed easy to show that as ψ_o falls our simple theory will also generate a burst in productivity growth along the transition (followed by a potential productivity slowdown). The reason for this burst in growth along the transition is twofold: (i) The general equilibrium force that decreases the incentive to in the new steady state innovate due to stiffer competition as S_t increased only realizes over time. Hence on impact—as ψ_o decreases—the incentive to do R&D increases unambiguously for all firms and productivity growth will increase initially. (ii) In a static sense the new steady state with a higher $S^\star$ is more efficient compared to the old one because the high productivity firms operate a larger fraction of the lines. This static efficiency gain must be realized along the transition leading yet again to high growth along the transition.

3.15. Theoretical extensions

The baseline model we laid out here is kept very parsimonious to show the minimum ingredients need to speak to the empirical facts in Section 2.. However, this tractable model can be augmented and generalized in various ways in order to make the theory more quantitative and without changing the key mechanism at work. Here we elaborate on some potential extensions.

The binary process efficiency is imposed to keep the structure as simple as possible. It is straightforward to generalize it for instance to a continuous distribution with and upper and lower bound. Then in general the whole (stationary) distribution will matter for the steady state and a simple sufficient statistic like $S^\star$ will not exist anymore.

One could also allow for some transition matrix between the process efficiency levels. It is not important that this variable is here assumed to be permanent. What is however important is that there is some persistence in the $\varphi(j)$ differences.

Another generalization we have analyzed is to relax the assumption $\gamma > \Delta$. With $\gamma < \Delta$ high productivity firms are less likely to be replaced by creative destruction since they remain the leader even if a low productivity type innovated upon them in the quality space. This then leads to a more dispersed markup distribution even with just two type of process efficiency. For instance with $\gamma^2 > \Delta$, high productivity firms can have a markup factor in a given line of either γ, $\Delta\gamma$, or Δ/γ whereas the low productivity type firms can have a markup of γ or γ^2/Δ.

The quadratic functional form of the overhead function gives rise to this simple linear-quadratic dynamic programming problem with a closed form solution. This property is maintained by adding an additional linear effect of $n(j)$ on overhead cost. But more generally, the overhead function could easily be generalized to any convex function.

Some generalizations are like having a CES final output production function instead of the Cobb-Douglas or having general CRRA preferences are also straightforward. With a CES and a elasticity of substitution σ larger than one it may happen that the markup of the high productivity type firm facing a low type second-best firm is no longer constraint by the fringe and the firm just sets the optimal markup $\frac{\sigma}{\sigma-1}$. Then the labor share within high productivity firms will decrease less as ψ_o decreases.

Since all firms operate an interval of lines of measure $n(j)$ firms will not loose all lines at once and consequently there is no firm exit. We however also characterized a version of the model where there are additional "small" firms that operate only one line. Then, as the rate of creative destruction decreases with the productivity slowdown also gross firm exit and entry rates decrease.

We also derived a version of our model where firms can do endogenous variety expansion. Then, as the span of control increases more varieties are created. As a consequence the R&D expenditure are spread over a larger number of product lines which reinforces the productivity slowdown of our baseline theory.

The baseline model here abstracts from physical capital. It is however straightforward to include physical capital by assuming a Cobb-Douglas production function for the variable output. The model would then predict that the physical capital share declines together with the labor income share (and the profit share goes up).

Finally, we analyze a version of our model in which we allow mergers and acquisitions. This theory predicts increased M&A activity during the transition.

4. Calibration

We calibrate the steady state model to assess the quantitative importance of the overhead cost mechanism. We compare two steady state economies that differ only by the value of the overhead cost parameter ψ_o. We define the initial steady state as 1948–1995 and new steady state as 2006–2018. We calibrate the two steady states to match eight moments in the data: 1) percent decline in aggregate labor share; 2) within firm labor share change relative to the labor share in the initial steady state; 3) top 10% concentration in over 1987–1992; 4) top 10% concentration over 2007–2012; 5) percentage change in top 10% concentration; 6) productivity growth rate over 1948–1995; 7) real interest rate in the initial steady state; and 9) R&D share of output in the initial steady state. We match the initial steady state growth rate exactly but give equal weights to all other moments. We compare the change in productivity growth rate in two steady state with the decline in the data, which is not targeted.

Table 3 displays the calibrated parameters. First, the concentration level is sensitive to the share of high productivity firms ϕ. All else equal, if ϕ is close to 1, the top 10% share is close to 10%. Lower ϕ with a sufficiently high Δ help to match the top 10% concentration in the data. We find that about 14% of the firms have high productivity. They enjoy an oversized market share because they are more than twice as efficient at the low productivity firms. Next, the quality step γ is set to match the growth rate in the baseline steady state given the rate of creative destruction, which is a function of other parameters. We calibrate it to close to 3. Given the growth rate of the economy, the real interest rate increases with the discount factor β. We calibrate β to 0.97. ψ_o is a normalization in the model. So we normalize the baseline value to 1. Holding other parameters fixed, a larger change in ψ_o generates larger change in labor share and concentration. We find that the model asks for 5% decline in the overhead costs parameter. Finally, we calibrate ψ_c using the R&D share.

Table 3: Baseline Parameter Values

Definition	Parameter	Value
1. share of H-type firms	ϕ	0.137
2. productivity gap	Δ	2.202
3. quality step	γ	2.935
4. discount factor	β	0.971
5. initial overhead cost	ψ_o	1
6. new, lower overhead cost	ψ_o'	0.954
7. R&D cost	ψ_c	5.816

Table 4 shows that we match the targets very well. Under these calibrated parameters, aggregate productivity growth declines from 1.8% to 1.4%, which is 60% of the actual decline. In Table 4, we also compare the aggregate labor share in the model with the data. Note that the labor share in the model is quite a bit lower than the measured labor share. This means that while we closely match the percentage decline in labor share, we under fit the percentage point decline in labor share. One way to see this is that our model generates one and half percentage points decline in labor share from rising aggregate markup. This leaves room for non-markup explanations such as technology and measurement for the decline in labor share.

Table 4: Baseline calibration

Targeted	Target	Model
1. % decline in labor share	4.4	4.2
2. within change/baseline labor share	0.075	0.077
3. top 10% concentration 1987–1992	0.675	0.675
4. top 10% concentration 2007–2012	0.728	0.728
5. change in top 10% concentration	7.9	7.8
6. productivity growth 1948–1995	0.018	0.018
7. baseline real interest rate	0.050	0.049
8. baseline R&D/PY	0.100	0.098
Untargeted	Data	Model
9. productivity growth 2006–2017	0.011	0.014
Share of growth decline explained		**60**%
10. aggregate labor share	0.541	0.348

Source: 1 and 10: BLS Employee only compensation share of value added, 2–5: Autor et al. (2017), 6 and 9: BLS MFP series

To clarify the mechanism in our model, Table 5 displays values for selected endogenous variables in the initial and new steady state. The decline in overhead costs encourages productive firms to expand, reducing the size of low productivity firms (lower $n^\star_L$) and increasing the size of high productivity firms (higher $n^\star_H$). Since the number of firms of each type stays constant, this translates into a higher fraction of products produced by the high productivity firms (higher $S^\star$). With the rise in $S^\star$, within firm labor share *rises* because firms are more likely to produce a product where the next best producer is a high productivity producer. The rise in within firm labor share means firms have lower markups. Similarly, the expect markup from innovating is also lower as firms are more likely to innovate on a product produced by a high productivity producer. As a result, firms reduce their R&D expenditures (lower

Z/Y), leading to a lower rate of creative destruction in the equilibrium (lower $z^{\star}$) and hence lower growth.

Table 5: Initial vs. new steady state

	Initial	New
1. creative destruction rate ($z^{\star}$)	0.017	0.013
2. share of H-type products ($S^{\star}$)	0.928	0.999
3. products per H-type firm ($n_H^{\star}$)	1	1.078
4. products per L-type firm ($n_L^{\star}$)	0.012	0.000
5. labor share of H-type firms	0.327	0.341
6. labor share of L-type firms	0.721	0.750
7. aggregate labor share	0.356	0.341
8. R&D/PY	0.098	0.073
9. total overhead/PY	0.185	0.205
10. rents/PY	0.362	0.381

Aggregate labor share declines in the new steady state. Since output is the sum of labor compensation, R&D investments, overhead costs and rents, we can decompose the decline in aggregate labor share into three parts: a change in R&D share, a change in overhead cost share and a change in rents share. In the new steady state, R&D investment is lower because there is less creative destruction. This pushes up the aggregate labor share. However, overhead costs rise because firms are bigger and rents share increases as more production is carried out by high markup firms. These forces dominates the decline in R&D share, resulting in the decline of aggregate labor share. The rise in overhead cost is consistent with the hypothesis proposed by Autor, Dorn, Katz, Patterson and Van Reenen (2017). It is similar in magnitude to the rise in the rents share. Despite the rise in aggregate rents share, innovation declines because the rise in the aggregate is due to reallocation and not an increase in firm-level markup.

Finally, recall that job reallocation across firms and establishments is trending down in the data, as shown in Figure ??. How might our model speak to this? Job reallocation across firms occurs when a firm's employment level rises (gross job creation) or falls (gross job destruction). In the data, this reallocation is partially due to firm entry and exit, which our baseline model does not have. But a significant component of job reallocation in the data is across surviving firms. In our model, firms add and subtract products from their portfolio due to creative destruction by themselves and their competitors. For simplicity our firms have a continuum of products, so this should ebb and flow nets out in steady state. But it is a short leap to a model in which firms have a finite number of products so that their employment levels rise and fall. See Garcia-Macia et al. (2018) for just such an analysis. Our model may speak more directly to job reallocation across establishments, if one makes the strong assumption that each plant is associated with given product line produced by the firm. Then plant entry and exit in the data can be compared to the rate of creative destruction in our model. As our model features falling long run growth, it implies falling long run job reallocation associated with product turnover.

5. Conclusion

We provide a new theoretical framework that can potentially account for a significant portion of the U.S. growth experience over the past 30 years: (i) a decline in the labor income share (driven by resource reallocation across firms as opposed to a declining labor income share within firms), (ii) a productivity slowdown (after a burst in productivity growth), (iii) rising concentration at the national level; and (iv) falling job reallocation rates. We argue that a significant part of these phenomena can be explained by ICT improvements in the mid-1990s to mid-2000s which increased the optimal boundary of (especially) the most efficient firms. In our theory, these firms enjoy higher markups; when they expand their reach into more markets, they raise average markups and lower the aggregate labor share. They expand by innovating on more product lines, bringing a temporary surge of growth. Within-firm markups eventually fall for both high and low productivity firms, as they are more likely to face high productivity competitors. This force ultimately drags down innovation, growth, and job reallocation.

We compared the steady states before and after a fall in the overhead cost for managing multiple product lines. In the next revision we plan to analyze the transition from one steady state to another, to quantify the magnitude and duration of the temporary gains in productivity growth. In the U.S. data, a 10-year period of rapid growth from 1996 to 2005 raised aggregate productivity about 10% above the pre-1996 trend. We may try to discipline the decline in overhead costs using the drop in ICT prices and the greater propensity of large firms to invest in ICT. We can also evaluate the welfare effects of falling overhead costs using the equivalent permanent change in consumption.

Another natural next step is to explore the cross-industry predictions of our theory and see if they hold up in the data. In particular, we might look at whether more intensively IT-using industries experienced bigger increases in concentration (paired with declining labor share, and a more pronounced boom-bust cycle of productivity growth. ICT can be used for more than just overhead, so we will try to gauge the overhead component of ICT by the difference in ICT investment rates of large versus small firms within industries.

Lastly, we are keen on exploring tax and subsidy policies in our quantitative framework. The decentralized equilibrium is suboptimal in due to markup dispersion across products as well as knowledge spillovers across firms (quality innovations build on previous innovations by other firms). It is possible that falling overhead costs reduce welfare, and that a welfare-improving policy response might be to (counterintuitively!) constrain the expansion of the most efficient firms. Conversely, we may find that the temporary surge in productivity more than justifies the lower long run growth prospects.

References

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

Akcigit, Ufuk and Sina Ates, "Declining Business Dynamism: Some Lessons from Modern Growth Theory and Patent Data," 2018.

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," Working Paper 23396, National Bureau of Economic Research May 2017.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium.," Working Paper 24007, National Bureau of Economic Research November 2017.

Barkai, Simcha, "Declining labor and capital shares," New Working Paper Series 2, Stigler Center for the Study of the Economy and the State 2016.

Bauer, Arthur, Jocelyn Boussard, and Daniel Laskhari, "Information Technology and Returns to Scale," Technical Report, Boston College, working paper 2018.

Bils, Mark, Peter J Klenow, and Benjamin A Malin, "Resurrecting the role of the product market wedge in recessions," *American Economic Review*, 2018, *108* (4-5), 1118–46.

Cao, Dan, Toshihiko Mukoyama, and Erick Sager, "Firm Growth through New Establishments," Technical Report, working paper 2018.

Chatterjee, Satyajit and Burcu Eyigungor, "Incumbency Disadvantage in U.S. National Policits: The Role of Policy Inertia and prospective Voting," Working paper 17-43 2017.

Crouzet, Nicolas and Janice Eberly, "Understanding weak capital investment: The role of market concentration and intangibles," in "prepared for the Jackson Hole Economic Policy Symposium" 2018.

David, Joel M and Venky Venkateswaran, "The sources of capital misallocation," *American Economic Review*, forthcoming.

De Loecker, Jan and Jan Eeckhout, "The Rise of Market Power and the Macroeconomic Implications," Working Paper 23687, National Bureau of Economic Research August 2017.

Eggertsson, Gauti B, Neil R Mehrotra, Sanjay R Singh, and Lawrence H Summers, “A contagious malady? Open economy dimensions of secular stagnation,” *IMF Economic Review*, 2016, *64* (4), 581–634.

Farhi, Emmanuel and François Gourio, “Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia,” Technical Report 25282, National Bureau of Economic Research 2018.

Fernald, John G, “Productivity and Potential Output before, during, and after the Great Recession,” *NBER macroeconomics annual*, 2015, *29* (1), 1–51.

_ , Robert E Hall, James H Stock, and Mark W Watson, “The Disappointing Recovery of Output after 2009,” *Brookings Papers on Economic Activity*, 2017, pp. 1–81.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow, “How Destructive is Innovation?,” NBER Working Papers 22953, National Bureau of Economic Research December 2018.

Gutiérrez, Germán and Thomas Philippon, “Investment-less growth: An empirical investigation,” Technical Report, National Bureau of Economic Research 2016.

_ and _ , “Declining Competition and Investment in the US,” Technical Report, National Bureau of Economic Research 2017.

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

Karabarbounis, Loukas, “The labor wedge: MRS vs. MPN,” *Review of Economic Dynamics*, 2014, *17* (2), 206–223.

_ and Brent Neiman, “The global decline of the labor share,” *The Quarterly Journal of Economics*, 2013, *129* (1), 61–103.

_ and _ , “Accounting for Factorless Income,” Working Paper 24404, National Bureau of Economic Research March 2018.

Klette, Tor Jakob and Samuel S. Kortum, “Innovating Firms and Aggregate Innovation,” *Journal of Political Economy*, October 2004, *112* (5), 986–1018.

Koh, Dongya, Raül Santaeulalia-Llopis, and Yu Zheng, “Labor share decline and intellectual property products capital,” *Available at SSRN 2546974*, 2016.

Liu, Ernest, Atif Mian, and Amir Sufi, "Low Interest Rates, Market Power, and Productivity Growth," Working Paper 25505, National Bureau of Economic Research January 2019.

Martinez, Joseba, "Automation, Growth and Factor Shares," 2017. Unpublished Working Paper, Department of Economics, New York University.

Rinz, Kevin, "Labor market concentration, earnings inequality, and earnings mobility," Technical Report, Center for Economic Studies, US Census Bureau 2018.

Rognlie, Matthew, "Deciphering the fall and rise in the net capital share: accumulation or scarcity?," *Brookings papers on economic activity*, 2016, *2015* (1), 1–69.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging trends in national and local concentration," Technical Report, National Bureau of Economic Research 2018.

Syverson, Chad, "Challenges to mismeasurement explanations for the US productivity slowdown," *Journal of Economic Perspectives*, 2017, *31* (2), 165–86.

Traina, James, "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements," New Working Paper Series 17, Stigler Center 2018.