Recent changes in firm dynamics and the nature of economic growth*

Markus Kondziella†
IIES, Stockholm University

Job Market Paper

This version: December 27, 2022. Latest version here

Abstract

In line with the US economy, market concentration and dispersion of revenue productivity, a popular measure of misallocation, increased in Sweden from 1997–2017. I document a novel finding in administrative data that provides a new perspective on the rise in market power and misallocation: firm growth accelerated starting in the 1990s. I reconcile the trends in a dynamic framework. Firms grow in size by expanding into new product markets and increase markups within their product markets through R&D. Firm sales and markup growth give rise to market concentration and misallocation at the industry level. The model rationalizes the empirical trends by reducing the cost of within-product market R&D for the most efficient firms (with higher markups) and raising the costs of entering new product markets. Despite generating misallocation and lowering the *level* of aggregate output, the rise in within-product market R&D creates a positive *growth* effect that increases the long-run growth rate by 0.5pp. The positive growth effect is a silver lining to the rise in concentration and misallocation. An R&D subsidy that counteracts the rise in misallocation, in fact, harms economic growth by lowering R&D incentives.

*I thank Timo Boppart, Per Krusell and Joshua Weiss for their advice and guidance on this project. I further thank Gualtiero Azzalini, Niklas Engbom, José-Elías Gallegos, Jonas Gathen, Sampreet Goraya, Basile Grassi, Ida Kristine Haavi, John Hassler, Kieran Larkin, Kurt Mitman, José Montalban, Alessandra Peter, Michael Peters and seminar and conference participants at the European Winter Meeting of the Econometric Society 2022, the Enter Jamboree 2022, the IIES, the Nordic Summer Symposium in Macroeconomics 2022, the Swedish Conference in Economics, the Swedish House of Finance, the Stockholm School of Economics and Stockholm University for their helpful comments.

†Institute for International Economic Studies (IIES), Stockholm University. E-mail: markus.kondziella@iies.su.se

1 Introduction

The US economy has experienced several trends over the last few decades. Sales concentration has been on the rise, dispersion of revenue productivity has increased, firm entry has slowed down, and the aggregate labor income share has declined. The trends have been associated with rising market power and misallocation (Akcigit and Ates, 2021; Decker, Haltiwanger, Jarmin and Miranda, 2017).[1] Even though the US economy has received a particularly large interest,[2] similar trends are observed globally.[3] Using high-quality administrative data on the universe of Swedish firms, I show that, in line with the US economy, sales concentration and revenue productivity dispersion increased during 1997–2017. While the trends at the industry level raise some cause for concern, I document a novel result at the firm level that sheds new light on the rise in market concentration and misallocation. Sales, revenue productivity, and employment growth have accelerated for firms established after 1997. The increase in growth rates translates into an upward shift of the firm's life cycle growth profile. This is the first study that documents an acceleration of firm life cycle growth starting in the 1990s.

To identify mechanisms that reconcile the empirical trends, I build a dynamic model where market concentration, misallocation, and firm growth arise endogenously. Firms compete through R&D for market shares and product markups. This generates sales and markup growth at the firm level and sales concentration and misallocation at the macro level. Competition through R&D links firm dynamics to economic growth.

Firm size and markup growth are the driving forces behind market concentration and misallocation. Firms grow in size by expanding into new product markets as in Klette and Kortum (2004). To enter a new product market, firms improve the product of the previous incumbent through R&D (creative destruction). Firms constantly compete for product markets through creative destruction. As a result, market concentration rises if firms expand into new product markets more aggressively. On the other hand, firms increase markups through R&D in their own product markets (own-innovation). Firms distance themselves through own-innovation from their competitors that produce the same product, allowing them to charge higher markups. Cross-sectional markup dispersion generates static output losses such that own-innovation is a source of misallocation. While both types of innovations (creative destruction and own-innovation) generate economic growth, they have different implications for firm sales and markup growth, market concentration, and misallocation. Not all firms grow in size and increase their markups at the same rate. Differences in permanent productivity levels lead to firms optimally choosing heterogeneous own-innovation and creative destruction efforts. More productive firms systematically increase their markups and sales faster than less productive firms in the economy.

[1]See Bils, Klenow and Ruane (2021) for a discussion about the limitations of revenue productivity dispersion as a measure of misallocation.

[2]For rising sales concentration and revenue productivity dispersion see Andrews, Criscuolo and Gal (2015); Autor, Dorn, Katz, Patterson and Van Reenen (2017); Decker, Haltiwanger, Jarmin and Miranda (2020); Grullon, Larkin and Michaely (2019); Van Reenen (2018). The decline in firm entry is documented in Decker, Haltiwanger, Jarmin and Miranda (2016); Gourio, Messer and Siemer (2014); Karahan, Pugsley and Şahin (2019). For the decline in the aggregate labor income share see Autor, Dorn, Katz, Patterson and Van Reenen (2020); Elsby, Hobijn and Şahin (2013); Karabarbounis and Neiman (2014); Kehrig and Vincent (2021); Lawrence (2015).

[3]See Andrews, Criscuolo and Gal (2016); Autor, Dorn, Katz, Patterson and Van Reenen (2020); Karabarbounis and Neiman (2014)

I take the model to the data by estimating unobserved model parameters, e.g., the cost of own-innovation and creative destruction, using Swedish administrative data. The model parameters are informed by firm-level moments. In particular, the estimated model parameters match firm life cycle growth in the model with the heterogeneous life cycle profiles that I observe in administrative data. To explain the increase in revenue productivity dispersion, sales concentration, and revenue productivity life cycle growth, I adopt a general approach that allows for any combination of parameter changes. To rationalize the empirical trends, the model asks for changes in innovation costs. A 31% decline in own-innovation costs for firms with a high productivity level coupled with a 13% increase in creative destruction costs for incumbents (11% for entrants) replicate the empirical trends quantitatively. The increase in creative destruction costs and the decrease in own-innovation costs capture the general notion that entering new markets has become more difficult, while firms find it easier to improve their own products in markets where they have already established themselves as the market leader.

The changes in own-innovation and creative destruction costs affect the long-run economy as follows: the fall in the own-innovation costs for firms with a high productivity level results in an acceleration of their markup life cycle growth. The acceleration in markup growth increases cross-sectional markup dispersion, generating misallocation. At the same time, faster markup accumulation incentivizes high-productivity firms to expand into new product markets. Together with a fall in firm entry due to rising creative destruction costs for entrants, this leads to sales concentration. The rise in markup dispersion introduces a permanent 0.7% loss in aggregate output. In contrast, the reallocation of market shares towards more efficient firms increases aggregate output permanently by 0.1% through a rise in aggregate productivity. Taken together, this results in a 0.6% permanent decline in the level of aggregate output. There is a positive side effect to the rise in misallocation and market concentration. Markup accumulation is the result of product-improving innovations. The rise in within-product market R&D increases the economy's long-run growth rate by almost half a percentage point. The positive growth effect on aggregate output works in the opposite direction as the negative level effect that arises through elevated markup dispersion. The positive growth effect puts the rise in misallocation in a more favorable light.

Creative destruction reduces misallocation as it lowers markup dispersion. I quantify the effects of a creative destruction subsidy on static efficiency and economic growth to assess the effectiveness of such a policy in addressing the rise in misallocation. For the policy counterfactuals, I introduce a government that subsidizes the firms' creative destruction expenditures through a profit tax. I find that a subsidy that lowers the creative destruction costs of a firm improves static efficiency but decreases the aggregate growth rate. Firms react to the increase in creative destruction by lowering their own-innovation efforts. Own-innovation is the primary source of economic growth, so the fall in own-innovation weighs heavily on the aggregate growth rate. The negative growth effect of the creative destruction subsidy echoes the well-known result that a high threat of replacement disincentivizes firms to engage in innovation. In my estimated model, the decline in own-innovation is significant enough to lower the aggregate growth rate. A subsidy on own-innovation costs instead successfully generates economic growth. The downside of such a subsidy is that it further worsens misallocation. An own-innovation subsidy that is financed by a 10% profit tax results in a permanent output loss of 1% relative to a subsidy on creative destruction. On the flip side, the aggregate growth rate with the own-innovation subsidy exceeds the growth rate with the creative destruction subsidy by 0.6pp. A subsidy that does not differentiate between the innovation types subsidizing total innovation expenditures of firms has a muted yet positive effect on economic growth.

This paper relates to the literature that documents trends in revenue productivity dispersion and sales concentration. Andrews, Criscuolo and Gal (2016), Decker, Haltiwanger, Jarmin and Miranda (2020), and Van Reenen (2018) find evidence for rising revenue productivity dispersion in the US and selected OECD countries. Autor, Dorn, Katz, Patterson and Van Reenen (2017), Grullon, Larkin and Michaely (2019) and Akcigit and Ates (2021) document rising sales concentration in the US. I confirm the trends of increased revenue productivity and sales concentration for the Swedish economy. Two advantages of my data are its quality and coverage. Balance sheet information comes directly from tax registries covering the universe of Swedish firms. Due to the comprehensiveness of the data, I show that the trends hold for the same set of firms. I further contribute a new finding to the literature. Firm life cycle growth of sales, employment, and revenue productivity accelerated for new firms established after 1997.

The proposed theory is based on Schumpeterian models of endogenous growth in the spirit of Aghion and Howitt (1992), Grossman and Helpman (1991), and Klette and Kortum (2004). These models are analytically tractable yet yield a realistic description of firm dynamics and capture salient features of the data (Lentz and Mortensen, 2008; Akcigit and Kerr, 2018). The framework in this paper particularly relates to Peters (2020) and Aghion, Bergeaud, Boppart, Klenow and Li (2019). Peters (2020) builds a framework of creative destruction and own-innovation.[4] In Peters (2020), all firms make identical innovation decisions as there is no ex-ante heterogeneity across firms. Ex-post shocks drive differences in realized markup, sales, and employment life cycle growth across firms. Sterk, Sedláček and Pugsley (2021) emphasize that ex-ante heterogeneity, rather than ex-post shocks, explains the differences in firm growth rates. I introduce heterogeneity in the permanent productivity level of firms, which gives rise to systematically different innovation and hence growth rates across firms. This heterogeneity is essential to explain the empirical trends and generates a cross-sectional firm size distribution that matches the data. Aghion, Bergeaud, Boppart, Klenow and Li (2019) feature heterogeneity in the permanent productivity level of firms. More productive firms charge higher markups and endogenously choose to operate more product lines. There is no incumbent own-innovation and firm entry such that this model abstracts from the firm life cycle.

The model application in this paper relates to the literature explaining recent trends in the US economy. Proposed drivers for the trends are an increasing importance of intangible capital and information and communications technology (ICT) (Aghion, Bergeaud, Boppart, Klenow and Li, 2019; Crouzet and Eberly, 2019; Chiavari and Goraya, 2020; De Ridder, 2019; Hsieh and Rossi-Hansberg, 2019; Weiss, 2019), demographic changes (Bornstein, 2018; Engbom, 2020; Hopenhayn, Neira and Singhania, 2018; Karahan, Pugsley and Şahin, 2019; Peters and Walsh, 2021), declining interest rates (Chatterjee and Eyigungor, 2019; Liu, Mian and Sufi, 2022), changes in the quality of ideas (Olmstead-Rumsey, 2019) or declining imitation rates (Akcigit and Ates, 2019). While matching trends at the industry level that the above papers focus on, I further explain the acceleration of firm life cycle growth. My analysis identifies a rise in own-innovation as a new driver behind the trends at the industry and firm level.

My theory relates to the literature on misallocation that has studied the effects of firm-level distor-

[4]Similarly, Akcigit and Kerr (2018) features a quality ladder model of endogenous growth with creative destruction and own-innovation.

tions on aggregate productivity. In Hsieh and Klenow (2009) and Restuccia and Rogerson (2008), misallocation arises from exogenous firm-specific wedges that are correlated with the firm's productivity. In Edmond, Midrigan and Xu (2018), firm productivity and markups correlate positively with size, whereas in Peters (2020), productivity and markups increase in firm age. My model captures misallocation from both the age and size dimensions separately. Firms increase their markups as they age, and firms with initially high markups expand faster in size.

The paper proceeds as follows. Section 2 introduces the data, section 3 documents the trends in the Swedish economy, and section 4 contains the theory. In section 5, I apply the model to the data to explain the empirical trends and quantify the implications for economic growth. The paper proceeds with policy counterfactuals that highlight a tradeoff between misallocation and economic growth. Section 6 concludes the paper.

2 Data

2.1 Data description

For the empirical results I link three data sets provided by Statistics Sweden (SCB). The first is *Företagens Ekonomi* (FEK), which covers information from balance sheets and profit and loss statements for the universe of Swedish firms. The unit of observation is the legal unit at annual frequency covering the period 1997-2017. Before 1997 FEK was a sample covering large Swedish firms. The data further contains information on the legal type and the industry code of the firm. I restrict myself to firms in the private economy with at least five employees. Using the detailed 5-digit industry codes, unless mentioned otherwise, all of the empirical analysis is conducted within industries.

The second data set is the *Producentprisindex* micro data, covering producer prices that enter the national producer price index (PPI). This data contains prices of goods for a representative sample of firms with at least 10 Mio. SEK net sales. In terms of employment, firms in the PPI data cover 22% of the cleaned FEK data in 2017 (28% of value added).

The auxiliary data set *Registerbaserad Arbetsmarknadsstatistik* (RAMS) provides additional information. It contains matched employer-employee data at the establishment level. I use it to define the universe of active employers and to get a measure of employment that is independent of the employment measure in the balance sheet data (measured in full-time equivalents).

Throughout the paper nominal variables are deflated to 2017-SEK values using the GDP deflator. The empirical analysis is carried out within 5-digit industry-years for which any finer deflation of variables (up to the 5-digit industry level if available) would have no effect.

2.2 Revenue productivity

Physical productivity measures a firm's technology, mapping inputs of production into physical output quantities. As large-scale data sets on the balance sheets of firms became available, the approach of measuring firm "productivity" using balance sheet variables (e.g. value added or sales) gained attraction. Productivity measures using value added or sales as output measures instead of physical output quantities has been labelled *revenue productivity*.

2.2.1 Measurement

One standard measures of revenue productivity is labor productivity, defined as

$$\text{Labor productivity}_{fst} \equiv \frac{VA_{fst}}{wL_{fst}},$$

where VA denotes value added and wL the wage bill for firm f in industry s. To account for the capital deepening of the firm Hsieh and Klenow (2009) introduced the concept of revenue TFP ("$TFPR$") as a measure of revenue productivity

$$TFPR_{fst} \equiv \frac{VA_{fst}}{K_{fst}^{\alpha_{st}}(wL_{fst})^{1-\alpha_{st}}}, \tag{1}$$

where K denotes capital and α the capital income share. I use $TFPR$ as my baseline measure of revenue productivity. All empirical results are robust to using labor productivity as the revenue productivity measure.[5] For the $TFPR$ computation I use fixed assets less depreciation as a measure of the firm's capital stock. Labor income shares, $(1-\alpha_{st})$, are computed as labor cost shares at the industry level, $\frac{wL_{st}}{wL_{st}+R\cdot K_{st}}$, where R is the rental price of capital set to 10% as in Hsieh and Klenow (2009).

2.2.2 Revenue productivity, misallocation and markups

The crucial distinction between physical productivity and revenue productivity is that revenue productivity, using the dollar value to measure output, captures both the firm's technology and the output price, whereas physical productivity captures just the firm's technology. The output price in return is affected by factors related to, e.g., industry competition or product quality. A firm that faces low competition or produces a product that is of higher quality than the product of its competitors potentially charges a relatively higher price. Low competition or high product quality therefore translate into high revenue productivity of the firm. In the model in section 4 firms innovate upon the quality of their products, which gives rise to life cycle growth in revenue productivity.

The fact that revenue productivity captures both the firm's technology and the output price is not a disadvantage relative to the physical productivity measure. One advantage of revenue productivity over physical productivity is that it allows for a comparison across firms. The technology of two firms that produce different output goods cannot be compared directly. Using the dollar value of output instead of physical quantities, revenue productivity can easily be compared across firms that produce different products. Further, by capturing the firm's output price, revenue productivity contains information that we can use to study misallocation and markups. In particular, dispersion of revenue productivity across firms is a popular measure of misallocation. The intuition is that firm-specific distortions that correlate with firm productivity (or quality) lead to aggregate output losses. Such distortions can take on the form of input or output taxes that correlate with the firm's productivity, or arise endogenously as in settings where more productive firms choose to constrain their production to charge higher markups. The point is that such distortions are reflected in the marginal revenue productivity of firms as profit maximizing firms equate the marginal revenue product with marginal costs. Dispersion of revenue productivity across firms is therefore a poplar indicator of such distortions in the economy.

[5] The empirical findings are also robust to using sales instead of value added and introducing intermediate inputs as a third production factor to eq. (1).

Table 1: Summary statistics (1997-2017)

	High TFPR			*Low TFPR*		
	Mean	SD	Obs.	Mean	SD	Obs.
*Value Added**	61.5	615.9	117,436	22.4	199.0	1,113,247
*Net Sales**	268.0	2,589.6	117,436	76.0	786.2	1,113,247
Employees	46.9	320.5	117,436	30.3	211.5	1,113,247
wL*	20.4	170.2	117,436	11.5	84.4	1,113,247
K*	38.2	454.5	117,436	29.5	548.0	1,113,247
ln *TFPR*	0.97	0.54	117,436	0.54	0.46	1,113,247

Note: *: in mio 2017-SEK. K is fixed assets minus depreciation, wL labor compensation. Firms that rank in the top decile of the industry $TFPR$ distribution for five years during 1997-2017 are classified as high $TFPR$ firms.

Revenue productivity is further related to the price-cost markup. As shown in De Loecker and Warzynski (2012), cost minimization of a firm that is a price taker in the input market and that operates a general production function using a flexible production input leads to the following expression for markups μ (here applied for labor as an input)

$$\mu_{fst} = \text{Output elasticity labor}_{fst} \times \frac{VA_{fst}}{wL_{fst}}.$$

The markup is proportional to revenue productivity. Most balance sheet data cover value added and labor costs such that revenue productivity of the firm (VA/wL) is straightforward to measure. Since estimation of the output elasticity requires functional form assumptions about the production function and data on both physical input quantities and physical output quantities, the output elasticity is generally hard to identify. With various assumptions one can infer information about markups from revenue productivity without estimating the output elasticity. First, in models where output is produced with a technology that is linear in labor (as in section 4) the output elasticity of labor equals unity. In this case the level of revenue productivity identifies the level of markups. Alternatively, under the assumption that firms within industries have the same output elasticity, dispersion of log markups across firms is equal to dispersion of log revenue productivity. Dispersion of revenue productivity is hence a measure of markup dispersion. A third alternative is to assume that the output elasticity is fixed over time (but potentially varies across firms). In that case markup growth of the firm is equal to revenue productivity growth. Hence, under various assumptions, revenue productivity measures the level of markups, markup dispersion in the cross-section or markup growth at the firm level.

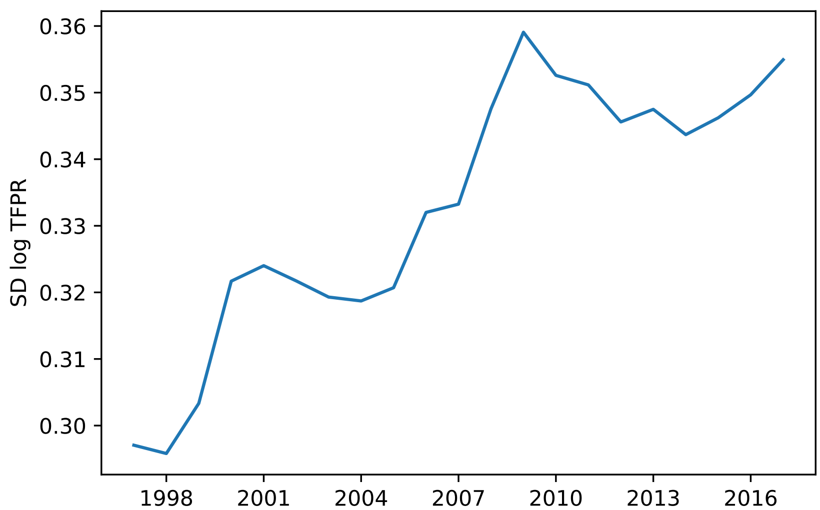
2.3 Descriptive statistics

Table 1 provides descriptive statistics for firms conditional on their revenue productivity. For this table, I classify firms into high and low revenue productivity firms according to their permanent revenue productivity level. A firm is labelled a high $TFPR$ firm if it ranks in the top decile of the industry $TFPR$ distribution for five years during 1997-2017. Table 1 shows that high $TFPR$ firms are on average larger than low $TFPR$ firms in terms of value added and sales, employ more workers (46.9 vs. 30.3), and own a larger capital stock. In this classification, high $TFPR$ firms have on average $e^{0.97-0.54} = 1.54$ higher $TFPR$ than low $TFPR$ firms. Note that those moments are pooled across all years and do not condition on the firm's age or industry.

3 Trends in the Swedish economy

Figure 1 shows the standard deviation of log revenue productivity ($TFPR$) within 5-digit industries during 1997–2017.[6] The standard deviation of log revenue productivity increased by 19.5% over this period. The increase in dispersion is particularly pronounced during 1997–2008, the years before the Great Recession. Despite the two recessions compressing revenue productivity dispersion temporarily a clear positive long-run trend is visible.

Figure 1: Rise in revenue productivity ($TFPR$) dispersion in Sweden



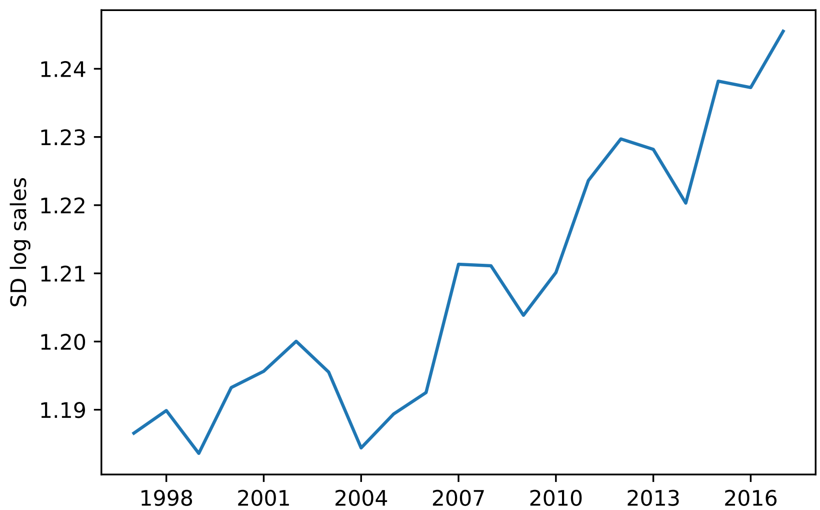
Notes: the figure shows the avg. standard deviation of log $TFPR$ within 5-digit industries with at least ten firms. $TFPR$ computed at the firm level as value added over a Cobb-Douglas composite of the capital stock and the wage bill (see section 2).

To compare magnitudes with the US, using Census data Decker, Haltiwanger, Jarmin and Miranda (2020) report an increase in the standard deviation of log labor productivity by roughly 4% during 1996–2013. Using their measure of revenue productivity, I observe an increase in the dispersion by 8% in Sweden for the same time period.[7] In Appendix A.1 I show that the rise in revenue productivity dispersion is robust to computing weighted standard deviations, alternative measures of revenue productivity or restricting to large industries. I further show that the increase in the standard deviation is not driven by a subset of firms but is due to a widening of the whole industry revenue productivity distribution.

[6] See section 2 for a detailed description of the data and variable construction.

[7] Industry classifications in the U.S. Census data are at the six-digit level and at the five-digit level in Swedish administrative data.

Figure 2: Rise in sales concentration in Sweden



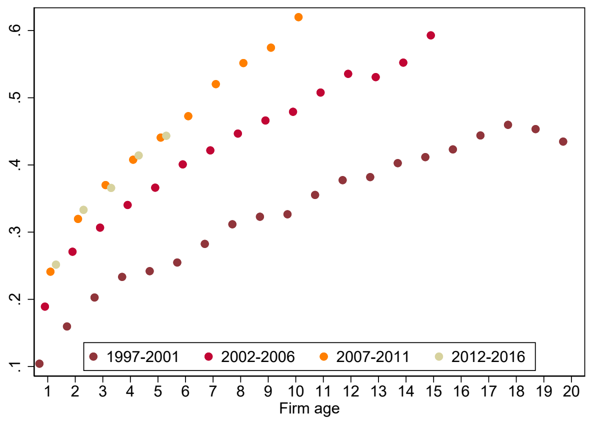
Notes: the figure shows the avg. standard deviation of log sales (shares) within 5-digit industries with at least ten firms.

The rise in revenue productivity dispersion is accompanied by an increase in sales concentration. Figure 2 shows the average industry sales concentration over time. I measure sales concentration within 5-digit industries using the standard deviation of log sales (or log sales shares). More dispersion in sales shares indicates higher sales concentration. The standard deviation of log sales increased by 5% from 1997 to 2017 providing evidence for a rise in sales concentration. As for revenue productivity dispersion, the early 2000s recession and the Great Recession compressed sales concentration temporarily, however the long-run trend is positive. The increase in sales concentration in Sweden is quantitatively similar to the one in the US. Sales shares of top-4 firms within 4-digit industries in U.S. Census data increased by roughly 4 percentage points (pp) during 1997–2012 (Akcigit and Ates, 2021; Autor, Dorn, Katz, Patterson and Van Reenen, 2020). Using this metric, sales concentration in Sweden equally increased by 4pp over the same time period (also measured at the 4-digit industry level). In Appendix A.2 I provide further robustness for the rise in sales concentration computing weighted standard deviations of sales shares or focusing on large industries.

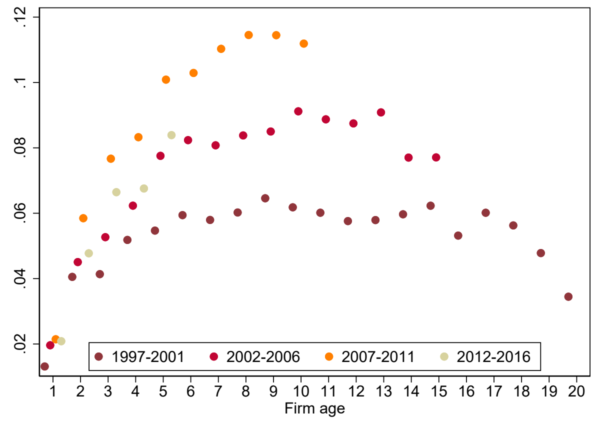
At the industry level, sales concentration and revenue productivity dispersion increased.[8] To understand the trends at the industry level, I shift the focus to the firm level. Figure 3 shows firms' life cycle growth of sales (left panel) and revenue productivity (right panel) measured as the log difference in sales and revenue productivity to age zero. To illustrate how life cycle growth of the firm has changed over time I show life cycle growth for different cohorts as indicated in the legend. The main observation is that starting in the early 2000s, growth rates of sales and revenue productivity have accelerated for new firms. This translates into steeper sales and revenue productivity life cycle gradients for younger cohorts. When looking at sales and revenue productivity growth during the

[8]Using the same data source as I do in this paper, Engbom (2020) further documents a decline in the firm entry rate in Sweden.

Figure 3: Firm life cycle growth (by cohort)



(a) Sales growth



(b) Revenue productivity growth

Notes: the figure shows life cycle growth of sales and revenue productivity measured as the log difference to age zero. Both measures are net of industry $\times$ year fixed effects. Cohorts pooled as indicated in the legend.

first four years of the firm, growth has increased by 78% (sales) and 25% (revenue productivity). A similar trend can be observed for employment life cycle growth, A.3. To obtain the life cycle growth of firms I regress log sales and log revenue productivity (net of industry $\times$ year fixed effects) on age dummies of the firm (age zero left out) and a constant. The age coefficients are displayed in the figure. In Appendix A.3 I provide robustness with respect to the cohort grouping. To the best of my knowledge this is the first paper that documents an acceleration in firm life cycle growth of sales, revenue productivity and employment.

In sum, at the same time as market concentration and misallocation increased at the industry level firm life cycle growth of sales, revenue productivity and employment accelerated. In the following section, I build a framework that speaks to the trends in market concentration, misallocation and firm growth jointly.

4 Model

4.1 Preferences and aggregate economy

The model is formulated in continuous time. The economy consists of a representative household that chooses the path of consumption C_t and wealth A_t to maximize lifetime utility

$$U = \int_0^\infty \exp(-\rho t) \ln C_t dt,$$

subject to the budget constraint $\dot{A}_t = r_t A_t + w_t L_t - C_t$ and a standard no-Ponzi game condition. ρ denotes the discount factor, r_t the interest rate and w_t the real wage. The household supplies one unit of labor inelastically such that $L_t = 1$. The optimality condition (Euler equation) for the household problem reads

$$\frac{\dot{C}_t}{C_t} = r_t - \rho.$$

Aggregate output is produced competitively using a Cobb-Douglas technology[9] over a continuum of different products indexed by i (time subscripts suppressed)

$$Y = \exp\left(\int_0^1 \ln\left[q_i y_i\right] di\right),$$

where y_i and q_i denote the quantity and quality of product i. Output is consumed entirely such that $Y = C$. Expenditure minimization leads to the standard demand function for product y_i

$$y_i = \frac{YP}{p_i}.$$

Here P is defined as the aggregate price index

$$P \equiv \exp\left(\int_0^1 \ln\left[p_i/q_i\right] di\right),$$

which is normalized to 1 (in each period).

4.2 Production

Firms can produce in every product market i with the following technology

$$y_{ij} = \varphi_j l_{ij},$$

where y_{ij} is the amount of product i produced by firm j, l_{ij} is the amount of labor hired, and φ_j denotes the physical productivity of firm j producing product i. A firm that is active in multiple markets produces the different products using the same productivity, i.e., φ_j varies with j, but not with i. Productivity of the firm is fixed over time, which captures the notion that some firms are persistently more efficient at producing than others. For simplicity, firms are either of the high or low productivity type, i.e., $\varphi_j \in \{\varphi^h, \varphi^l\}$ with $\varphi^h/\varphi^l > 1$.

In previous models of creative destruction and own-innovation, the notion of the firm is a collection of random product lines. With differences in productivity across firms, a firm is a collection of product lines, among which the firm produces with a technology that is potentially distinct from the technology of its competitors. Apart from changing the interpretation of the firm, permanent differences in productivity across firms have implications for firm dynamics. More productive firms will charge higher markups and therefore have a larger incentive to expand into new product lines: more productive firms endogenously grow faster in size. Without permanent productivity differences, firms choose identical innovation rates. In this case, firm size growth is purely determined by ex-post shocks. Sterk, Sedláček and Pugsley (2021) emphasize that ex-ante heterogeneity rather than ex-post shocks explain differences in firm growth rates. Productivity differences that lead to differences in the optimal innovation rates across firms introduce a deterministic component to firm growth.

[9]The Cobb-Douglas technology can be relaxed to a CES aggregator. In fact, with a CES aggregator own-innovation generates both misallocation and market concentration. Rising own-innovation (as highlighted later on) plays an even larger role in explaining the empirical trends in a CES setting than under the Cobb-Douglas technology.

4.3 Static allocation

Taking the distribution of product qualities and the number of firms as exogenous in this section, I characterize static allocations at the product, firm and aggregate level.

4.3.1 Product level

Firms within a product market i compete in prices (Bertrand competition). In equilibrium, only the firm with the highest quality adjusted productivity $q_{ij}\varphi_j$ produces product i.

Under Bertrand competition, the leader (the firm with the highest quality adjusted productivity) engages in limit pricing and sets its price equal to the quality-adjusted marginal costs of the follower (the firm with second highest quality adjusted productivity). The leader's price in equilibrium is hence given by

$$p_{ij} = \frac{q_{ij}}{q_{ij'}} \frac{w}{\varphi_{j'}}, \tag{2}$$

where j' indexes the follower in a market. According to eq. (2), the price that the leader sets is increasing in the quality gap between the leader and the follower.

The equilibrium price-cost markup in market i for producer j is defined as the output price over marginal costs, hence

$$\mu_{ij} \equiv \frac{p_{ij}}{w/\varphi_j} = \frac{q_{ij}}{q_{ij'}} \frac{\varphi_j}{\varphi_{j'}}. \tag{3}$$

The leader's markup for product i is increasing in the quality and productivity gap between the leader and the follower.

The price setting of the leader gives rise to the following equilibrium profits for product i

$$\pi_{ij} = p_{ij} y_{ij} - w l_{ij} = Y \left(1 - \frac{1}{\mu_{ij}}\right),$$

with labor demand for product i given by

$$l_{ij} = \frac{Y}{w} \mu_{ij}^{-1}.$$

Employment in product line i is decreasing in the markup.

To see the relationship between markups and revenue productivity, I define $TFPR_{ij}$ of the incumbent[10] in line i as

$$TFPR_{ij} \equiv \frac{p_{ij} y_{ij}}{w l_{ij}} = \frac{p_{ij}}{w} \varphi_j = \frac{q_{ij}}{q_{ij'}} \frac{\varphi_j}{\varphi_{j'}} = \mu_{ij}. \tag{4}$$

Revenue productivity in line i is equal to the markup, which results from an output elasticity of labor equal to unity. Eq. (4) highlights the distinction between revenue productivity and physical productivity. Physical productivity φ_j, measuring the firm's technology, enters revenue productivity, however revenue productivity further reflects the output price captured by the physical productivity of the second best firm and the quality gap.

[10] In the following I refer to leader and incumbent equivalently.

4.3.2 Firm level

Summing employment per product over the set of products that firm j is producing in, N_j, gives employment at the firm-level:

$$l_j = \sum_{i \in N_j} l_{ij} = \frac{Y}{w} \left(\sum_{i \in N_j} \mu_{ij}^{-1} \right).$$

Employment at the firm level is increasing in the number of products that the firm is producing.

At the firm level I define revenue productivity as firm sales over the firm's wage bill:

$$TFPR_j \equiv \frac{\sum_{i \in N_j} p_{ij} y_{ij}}{w \sum_{i \in N_j} l_{ij}} = \frac{\sum_{i \in N_j} p_{ij} y_{ij}}{w l_j} = \left(\frac{1}{n_j} \sum_{i \in N_j} \mu_{ij}^{-1} \right)^{-1}, \tag{5}$$

where n_j denotes the cardinality of N_j. $TFPR_j$ is a harmonic mean of the markups the firm charges for its products.

Sales of the firm are $n_j Y$, which follows from the fact that revenue per line is equalized.

4.3.3 Aggregate level

Summing firm employment over all firms yields the total work force for production:

$$L_P = \int_j l_j dj = \frac{Y}{w} \int_0^1 \mu_{ij}^{-1} di. \tag{6}$$

An expression for the wage can be found from the markup equation (3):

$$w = \frac{p_{ij}}{\mu_{ij}} \varphi_j = \frac{p_{ij}}{q_{ij}} \frac{q_{ij} \varphi_j}{\mu_{ij}}.$$

Taking logs and integrating over all products gives

$$w = \exp\left(\int_0^1 \ln \left[p_{ij}/q_{ij} di \right] \right) \times \exp\left(\int_0^1 \ln q_{ij} di \right) \times \exp\left(\int_0^1 \ln \varphi_{j(i)} di \right) \times \exp\left(\int_0^1 \ln \mu_{ij}^{-1} di \right).$$

As the final good is the numeraire, the wage expression simplifies to

$$w = \exp\left(\int_0^1 \ln q_{ij} di \right) \times \exp\left(\int_0^1 \ln \varphi_{j(i)} di \right) \times \exp\left(\int_0^1 \ln \mu_{ij}^{-1} di \right).$$

To find an expression for aggregate output rearrange equation (6):

$$Y = \frac{w}{\int_0^1 \mu_{ij}^{-1} di} L_P.$$

Inserting the wage expression gives

$$Y = Q \Phi \mathcal{M} L_P,$$

where

$$Q = \exp\left(\int_0^1 \ln q_{ij} di\right)$$
$$\Phi = \exp\left(\int_0^1 \ln \varphi_{j(i)} di\right)$$
$$\mathcal{M} = \frac{\exp\left(\int_0^1 \ln \mu_{ij}^{-1} di\right)}{\int_0^1 \mu_{ij}^{-1} di}$$

Aggregate output Y depends on geometric averages of quality Q and productivity Φ as well as on the dispersion of markups $\mathcal{M}$ and the total labor force L_P. Aggregate TFP is given by $Q\Phi\mathcal{M}$. The degree of markup dispersion is captured by $\mathcal{M}$. $\mathcal{M}$ is less (or equal) than unity as a geometric mean is weakly lower than an arithmetic mean. Any markup dispersion drives $\mathcal{M}$ below unity causing allocative inefficiencies. Only if markups are equalized, it holds that $\mathcal{M} = 1$.

Using again equation (6), monopoly power affects factor prices through a reduction in labor demand. The aggregate labor income share is given by

$$\Lambda \equiv \frac{wL_P}{Y} = \int_0^1 \mu_{ij}^{-1} di.$$

Aggregate TFP depends on the dispersion of markups. The aggregate labor income share depends on the level of markups.

Due to the Cobb-Douglas demand, the sales-weighted average markup equals the unweighted markup

$$E^{sales}[\mu] \equiv \int_0^1 \mu_i \frac{p_i y_i}{Y} di = \int_0^1 \mu_i di.$$

The cost-weighted average markup is equal to the inverse of the aggregate labor income share

$$E^{cost}[\mu] \equiv \int_0^1 \mu_i \frac{w\ell_i}{wL_p} di = \int_0^1 \mu_i \frac{\ell_i}{L_p} di = \frac{Y}{wL_P} = \Lambda^{-1}.$$

It is easy to show that the cost-weighted average markup is weakly lower than the sales-weighted one.[11] Intuitively, a high markup in a given line reduces the labor demand and hence the labor cost in that line.

4.4 Dynamic firm problem

Firms continuously improve the quality of products, q_i, in the economy through two different types of innovations. Own-innovation raises the quality of an item that the firm itself is producing, whereas through creative destruction the firm improves the quality of a competitor's product. Item quality is improved step-wise such that every time quality is improved (either through creative destruction or through own innovation) quality increases by a factor of λ. As Aghion, Bergeaud, Boppart, Klenow and Li (2019) I assume $\lambda > \varphi^h/\varphi^l$. This assumption guarantees that the firm with the highest quality version in a product line is always the active producer. Denote by λ^{Δ_i} the

[11] Using Jensen's inequality: $\frac{1}{\int_0^1 \mu_i di} \leq \int_0^1 \mu_i^{-1} di \Leftrightarrow E^{cost} \equiv \frac{1}{\int_0^1 \mu_i^{-1} di} \leq \int_0^1 \mu_i di \equiv E^{sales}$.

ratio of product qualities between the active producer and the second best firm (firm with second highest value of $q_{ij}\varphi_j$) in product line i such that

$$\lambda^{\Delta_i} = \frac{q_{ij}}{q_{ij'}}.$$

Firm profits are determined by the markups in each product line, which depend on the productivity and quality gap to the follower. To infer current profits of the firm the markups per line are sufficient, however for the dynamic problem of the firm one needs to keep track of the firm's productivity separately. The (expected) productivity gap in a new line depends on the productivity type of the firm. To save on notation denote by $[\mu_i]$ the set of markups in the firm's product lines. Firm profits are given by

$$\pi_t(n, [\mu_i]) = \sum_{k=1}^{n} Y_t \left(1 - \frac{1}{\mu_{kjt}}\right) = \sum_{k=1}^{n} Y_t \left(1 - \frac{1}{\lambda^{\Delta_{kt}}} \frac{1}{\frac{\varphi_{kj}}{\varphi_{kj'}}}\right) \equiv \sum_{k=1}^{n} \pi(\mu_{kt}).$$

Whereas $\pi_t(n, [\mu_i])$ denotes total firm profits, $\pi(\mu_{kt})$ denotes product line profits.

Firms choose the rate of own-innovation I_i and the rate of expansion x_i (own creative destruction) for its products. When choosing optimal own-innovation and expansion rates, firms take aggregate output Y_t, the real wage w_t, the share of lines operated by high productivity producers S_t, the interest rate r_t and the rate of creative destruction τ_t as given. Denoting the time derivative by $\dot{V}_t^h()$, the value function of a high process efficiency firm indexed by h satisfies the following HJB equation:

$$\begin{aligned}
r_t V_t^h(n, [\mu_i]) - \dot{V}_t^h(n, [\mu_i]) = \\
&\sum_{k=1}^{n} \underbrace{\pi(\mu_k)}_{\text{Flow profits}} + \sum_{k=1}^{n} \underbrace{\tau_t \Big[V_t^h\left(n-1, [\mu_i]_{i\neq k}\right) - V_t^h(n, [\mu_i])\Big]}_{\text{Agg. creative destruction}} \\
&+ \max_{[x_k, I_k]} \Bigg\{ \sum_{k=1}^{n} \underbrace{I_k \Big[V_t^h\left(n, [[\mu_i]_{i\neq k}, \mu_k \times \lambda]\right) - V_t^h(n, [\mu_i])\Big]}_{\text{Own-innovation}} \\
&+ \sum_{k=1}^{n} \underbrace{x_k \Big[S_t V_t^h\left(n+1, [[\mu_i], \lambda]\right) + (1 - S_t) V_t^h\left(n+1, \big[[\mu_i], \lambda \times \varphi^h/\varphi^l\big]\right) - V_t^h(n, [\mu_i])\Big]}_{\text{Product expansion}} \\
&- \underbrace{w_t \Gamma^h\left([x_i, I_i]; n, [\mu_i]\right)}_{\text{Research costs}} \Bigg\}.
\end{aligned}$$

The HJB equation of a low process efficiency firm follows the same expression with the product expansion term replaced by $x_k\Big[S_t V_t^l\left(n+1, \big[[\mu_i], \lambda \times \varphi^l/\varphi^h\big]\right) + (1 - S_t)V_t^l\left(n+1, [[\mu_i], \lambda]\right) - V_t^l(n, [\mu_i])\Big]$. As in Akcigit and Kerr (2018) and Peters (2020) the value of a firm consists of flow profits, research costs and three parts related to own-innovation, expansion and creative destruction. At rate τ_t (determined in equilibrium) the firm loses one of its n products in which case $n-1$ of its products remain. The firm chooses own-innovation rates I_k and expansion rates x_k for each product in which case the quality of one of the firm's products increases by one step (quality gap increases to $\Delta + 1$), or the firm acquires a new product (the firm produces $n + 1$ products).

The formulation of the value function differs from previous models in the literature that feature own-innovation and creative destruction along two dimensions. First, the value function (and the resulting firm policies) are specific to the productivity type of the firm. Second, the value function depends on the distribution of firm productivity types across product lines through the product expansion term. The expected value of acquiring a new product line depends on the probability of facing a high productivity incumbent S_t. Firms take S_t as given when making optimal decisions, however affect it through their decisions in equilibrium.

$\Gamma^h([x_i, I_i]; n, [\mu_i])$ is the productivity type specific cost of researching for both own-innovation and expansion. I assume the following flexible functional form

$$\Gamma^h([x_i, I_i]; n, [\mu_i]) = \sum_{k=1}^{n} c(x_k, I_k; \mu_k) = \sum_{k=1}^{n} \left[\mu_k^{-1} \frac{1}{\psi_I^h} (I_k)^\zeta + \frac{1}{\psi_x} (x_k)^\zeta \right],$$

with $\zeta > 1$. Profits within a product line are concave in the markup. The incentives to own-innovate therefore decrease with the quality gap that the firm has accumulated. I scale the own-innovation costs by the inverse markup, which renders the product line specific own-innovation rate to be independent of the product markup in equilibrium. I introduce a productivity type specific own-innovation cost shifter (ψ_I^h, ψ_I^l) to match markup life cycle growth specific to the productivity type of the firm in the data. The cost shifter for creative destruction ψ_x is the same for both productivity types. This does not, however, imply that all firms expand at equal rates in equilibrium. High productivity firms have a larger incentive to expand into new product markets as they charge higher markups upon entry. Heterogeneity in the cost shifter is therefore not necessary to generate productivity specific expansion rates x^h and x^l. All parameters are estimated by matching life cycle growth for high and low process efficiency firms.

Firm entry is determined as follows: using a linear production technology potential entrants produce a flow of marketable ideas ψ_z per unit of labor that improves on the quality of a randomly selected product line. Entrants start with a one-step quality gap. I assume that after entering, firms get assigned the high productivity type with probability p^h. Denoting by z_t the equilibrium flow rate of entry, the free entry condition reads

$$p^h \left(\underbrace{S_t V_t^h(1, \lambda) + (1 - S_t) V_t^h(1, \lambda \times \varphi^h/\varphi^l)}_{\text{Expected value of entering with } \varphi^h} \right) + (1 - p^h) \left(\underbrace{S_t V_t^l(1, \lambda \times \varphi^l/\varphi^h) + (1 - S_t) V_t^l(1, \lambda)}_{\text{Expected value of entering with } \varphi^l} \right) \leq \frac{1}{\psi_z} w_t,$$

which holds with equality if there is positive entry, i.e., $z_t > 0$. With high and low process efficiency firms expanding at different rates firm entry is necessary to generate a stable distribution over productivity types.

4.5 Distribution over quality and productivity gaps

In this section I characterize the two-dimensional distribution of incumbents' quality and productivity gaps across all product lines, $\nu_t\left(\Delta, \frac{\varphi_j}{\varphi_{j'}}\right)$. This two-dimensional distribution determines the markup distribution across all product lines, which is required for the solution of the model. From the firm's maximization problem it will turn out that along the balanced growth path the own-innovation and expansion rate are time invariant and do not depend on the quality gap in a product line, i.e., they only vary across the firm's productivity type. I characterize the distribution for the case where I and x only depend on the productivity type of the firm.

The distribution is characterized by a set of infinitely many differential equations. The change in the mass of product lines with a quality gap of λ^{Δ} and a productivity gap of $\frac{\varphi_j}{\varphi_{j'}}$ is

$$\dot{\nu}_t\left(\Delta, \frac{\varphi_j}{\varphi_{j'}}\right) = I^l \nu_t\left(\Delta - 1, \frac{\varphi_j}{\varphi_{j'}}\right) - \nu_t\left(\Delta, \frac{\varphi_j}{\varphi_{j'}}\right)(I^l + \tau_t) \quad \text{for} \quad \Delta \geq 2, \frac{\varphi_j}{\varphi_{j'}} \in \left\{\frac{\varphi^l}{\varphi^h}, \frac{\varphi^l}{\varphi^l}\right\}$$
$$\dot{\nu}_t\left(\Delta, \frac{\varphi_j}{\varphi_{j'}}\right) = I^h \nu_t\left(\Delta - 1, \frac{\varphi_j}{\varphi_{j'}}\right) - \nu_t\left(\Delta, \frac{\varphi_j}{\varphi_{j'}}\right)(I^h + \tau_t) \quad \text{for} \quad \Delta \geq 2, \frac{\varphi_j}{\varphi_{j'}} \in \left\{\frac{\varphi^h}{\varphi^h}, \frac{\varphi^h}{\varphi^l}\right\}. \tag{7}$$

For the case of a unitary quality gap, $\Delta = 1$:

$$\begin{aligned}
\dot{\nu}_t\left(1, \frac{\varphi^l}{\varphi^h}\right) &= (1 - S_t)x^l S_t + z_t(1 - p^h)S_t - \nu_t\left(1, \frac{\varphi^l}{\varphi^h}\right)(I^l + \tau_t) \\
\dot{\nu}_t\left(1, \frac{\varphi^l}{\varphi^l}\right) &= (1 - S_t)x^l(1 - S_t) + z_t(1 - p^h)(1 - S_t) - \nu_t\left(1, \frac{\varphi^l}{\varphi^l}\right)(I^l + \tau_t) \\
\dot{\nu}_t\left(1, \frac{\varphi^h}{\varphi^h}\right) &= S_t x^h S_t + z_t p^h S_t - \nu_t\left(1, \frac{\varphi^h}{\varphi^h}\right)(I^h + \tau_t) \\
\dot{\nu}_t\left(1, \frac{\varphi^h}{\varphi^l}\right) &= S_t x^h(1 - S_t) + z_t p^h(1 - S_t) - \nu_t\left(1, \frac{\varphi^h}{\varphi^l}\right)(I^h + \tau_t).
\end{aligned} \tag{8}$$

The law-of-motion for the mass of product lines with quality gap Δ and a given productivity gap is characterized by the inflow minus the outflow of mass. Outflows are due to own-innovation (quality gap increases from Δ to $\Delta + 1$) and creative destruction (quality gap gets reset from Δ to unity). Inflows are due to creative destruction (for $\Delta = 1$) or due to own-innovation (for $\Delta \geq 2$) as the quality gap increases from $\Delta - 1$ to Δ.

Using the measure over quality and productivity gaps, the aggregate share of products with a high productivity incumbent is defined as

$$S_t = \sum_{i=1}^{\infty} \left[\nu_t\left(i, \frac{\varphi^h}{\varphi^h}\right) + \nu_t\left(i, \frac{\varphi^h}{\varphi^l}\right)\right]. \tag{9}$$

The rate of creative destruction is then

$$\tau_t = z_t + x^h S_t + x^l(1 - S_t).$$

Note that τ_t is a function of S_t and S_t in return is a function of ν. The rate of creative destruction therefore depends on the distribution of productivity types. From eq. (9) it follows

$$\dot{S}_t = \sum_{i=1}^{\infty} \left[\dot{\nu}_t\left(i, \frac{\varphi^h}{\varphi^h}\right) + \dot{\nu}_t\left(i, \frac{\varphi^h}{\varphi^l}\right)\right] = S_t x^h + z_t p^h - S_t \tau_t = S_t(1 - S_t)(x^h - x^l) + z_t\left(p^h - S_t\right). \tag{10}$$

The third term has an intuitive interpretation. The change in the aggregate share of high productivity firms is due to high productivity firms replacing existing firms either through expansion or entry (first two terms) or because they themselves get replaced (third term). Note that in this definition high productivity firms replacing high productivity firms are included both in the terms $S_t x^h + z_t p^h$ and $S_t \tau_t$. Using the definition of τ and expanding one can show that

$$\dot{S}_t = S_t x^h (1 - S_t) - (1 - S_t) x^l S_t + z_t \left(p^h (1 - S_t) - (1 - p^h) S_t \right).$$

Changes in S_t are due to high productivity firms expanding into markets with low productivity incumbents (first term), low productivity firms expanding into markets with high productivity incumbents (second term), high productivity entrants replacing low productivity incumbents and low productivity entrants replacing high productivity incumbents (final term).

4.6 Balanced growth path (BGP) characterization

Labor market clearing requires that production labor L_{Pt} and research labor L_{Rt} add up to one, the aggregate labor endowment:

$$1 = L_{Pt} + L_{Rt}$$

A balanced growth path (BGP) is a set of allocations $[x_{it}, I_{it}, \ell_{it}, z_t, S_t, y_{it}, C_t]_{it}$ and prices $[r_t, w_t, p_{it}]_{it}$ such that firms choose $[x_{it}, I_{it}, p_{it}]$ optimally, the representative household maximizes utility choosing $[C_t, y_{it}]_{it}$, the growth rate of aggregate variables is constant, the free-entry condition holds, all markets clear and the cross sectional joint distribution of quality and productivity gaps ν_t is stationary.

In Appendix B.1 I show that the value function of the high productivity type firm takes on the following expression (with h replaced by l for the low productivity type)

$$\begin{aligned} V_t^h(n, [\mu_i]) &= V_{t,P}^h(n) + \sum_{k=1}^{n} V_{t,M}^h(\mu_k) \\ &= n \frac{1}{(\rho + \tau)} \frac{\zeta - 1}{\psi_x} (x^h)^\zeta w_t + \sum_{k=1}^{n} \frac{\pi^h(\mu_k) + \frac{\zeta - 1}{\psi_I^h} (I^h)^\zeta w_t \mu_k^{-1}}{\rho + \tau}. \end{aligned} \tag{11}$$

The value function is additive across products. The first part of the value function that represents the option value of expanding into new markets scales linearly in the number of products. The second part consists of flow profits and the option value to increase markups further. Both terms are scaled by the sum of the discount factor and the rate of creative destruction, the rate at which products get replaced. The value function is productivity type specific. First, the option value to expand is productivity type specific since profits upon entry depend on firm productivity. Second, firm productivity enters flow profits and the option value to increase markups in the future.

In the following I characterize the two dimensional distribution of quality and productivity gaps as a function of firm policies. This allows for optimal policies and the distribution to be solved jointly. The steady state distribution over quality and productivity gaps is found by setting the differential equations characterizing the law-of-motion in eq. (7) and (8) equal to zero. The stationary distribution is characterized by

$$\nu\left(\Delta,\frac{\varphi^l}{\varphi^h}\right)=\left(\frac{I^l}{I^l+\tau}\right)^{\Delta}\frac{(1-S)x^lS+z(1-p^h)S}{I^l}$$
$$\nu\left(\Delta,\frac{\varphi^l}{\varphi^l}\right)=\left(\frac{I^l}{I^l+\tau}\right)^{\Delta}\frac{(1-S)x^l(1-S)+z(1-p^h)(1-S)}{I^l}$$
$$\nu\left(\Delta,\frac{\varphi^h}{\varphi^h}\right)=\left(\frac{I^h}{I^h+\tau}\right)^{\Delta}\frac{Sx^hS+zp^hS}{I^h}$$
$$\nu\left(\Delta,\frac{\varphi^h}{\varphi^l}\right)=\left(\frac{I^h}{I^h+\tau}\right)^{\Delta}\frac{Sx^h(1-S)+zp^h(1-S)}{I^h}$$

Summing the measure over all quality gaps for a given productivity gap yields the share of products with a given productivity gap:

$$S_{\varphi^l,\varphi^h}\equiv\sum_{i=1}^{\infty}\nu\left(i,\frac{\varphi^l}{\varphi^h}\right)=\frac{(1-S)x^lS+z(1-p^h)S}{\tau}$$
$$S_{\varphi^l,\varphi^l}\equiv\sum_{i=1}^{\infty}\nu\left(i,\frac{\varphi^l}{\varphi^l}\right)=\frac{(1-S)x^l(1-S)+z(1-p^h)(1-S)}{\tau}$$
$$S_{\varphi^h,\varphi^h}\equiv\sum_{i=1}^{\infty}\nu\left(i,\frac{\varphi^h}{\varphi^h}\right)=\frac{Sx^hS+zp^hS}{\tau}$$
$$S_{\varphi^h,\varphi^l}\equiv\sum_{i=1}^{\infty}\nu\left(i,\frac{\varphi^h}{\varphi^l}\right)=\frac{Sx^h(1-S)+zp^h(1-S)}{\tau}$$

The share of products with a given productivity gap is in the BGP equal to the share of total creative destruction that starts with a quality gap of unity for that given productivity gap at each instant in time. From this, S is implicitly defined by

$$S=S_{\varphi^h,\varphi^h}+S_{\varphi^h,\varphi^l}=\frac{Sx^h+zp^h}{\tau}.$$

In Appendix B.3, I show that the share of product lines with a quality gap smaller than m conditional on the productivity gap follows

$$P\left(\lambda^{\Delta}\leq m,\frac{\varphi^l}{\varphi^h}\right)=S_{\varphi^l,\varphi^h}\left(1-m^{-\theta_l}\right),$$

where I denote $\theta_l=\frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}$. Conditional on the productivity gap, quality gaps follow a Pareto distribution with parameter θ. The Pareto shape parameter is affected by the rate of own-innovation I. The more own-innovation the more mass is in the tail of the quality gap distribution. In Peters (2020), markups follow a Pareto distribution. I introduce ex-ante differences in firm productivities, which leads to the result that markups *conditional* on the productivity gap between incumbents and second best firms are Pareto distributed.

The two-dimensional distribution of productivity and quality gaps across all product lines characterizes the markup distribution. The average markup in the economy is

$$E[\mu] = \sum_{k\in\{h,l\}} \underbrace{\frac{\theta_k}{\theta_k - 1}}_{\text{Churn}} \sum_{n\in\{h,l\}} \underbrace{\frac{\varphi_k}{\varphi_n}}_{\text{Prod. gap}} \times \underbrace{S_{\varphi_k,\varphi_n}}_{\text{Distribution over prod. gaps}} .$$

The derivation of the moments of the markup distribution can be found in the Appendix, section B.4. The average markup depends on the speed of own-innovation relative to creative destruction (Churn), the size of productivity gaps and third the distribution over productivity gaps.

The markup distribution further characterizes the aggregate labor income share Λ that enters the labor market clearing condition, see Appendix B.3.

$$\Lambda = \sum_{k\in\{h,l\}} \frac{\theta_k}{\theta_k + 1} \sum_{n\in\{h,l\}} \frac{1}{\varphi_k/\varphi_n} S_{\varphi_k,\varphi_n}. \tag{12}$$

Λ and S are solved jointly with the optimality conditions of the firm, the labor market clearing condition and the free entry condition. The equilibrium conditions are derived in Appendix B.2.

Growth in this model is the result of quality improvements. This occurs through either own-innovation, firm expansion or firm entry. The steady-state growth rate of aggregate variables is derived in Appendix B.5.

$$g = \frac{\dot{Q}_t}{Q_t} = \left(\underbrace{SI^h + (1-S)I^l}_{\text{Incumbent own-innovation}} + \underbrace{Sx^h + (1-S)x^l}_{\text{Incumbent product expansion}} + \underbrace{z}_{\text{Entry}} \right) \times \ln(\lambda) \tag{13}$$

The growth rate is equal to the aggregate arrival rate of innovation times the step size of innovation. Since firms with different productivities innovate with different rates, the growth rate depends on the distribution over productivity types.

The TFP misallocation measure $\mathcal{M}$ that captures the static loss in output that arises from markup dispersion is derived in Appendix B.3.

$$\mathcal{M} = \frac{e^{\sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \left[S_{\varphi_k,\varphi_n} \left(\ln\left(\frac{1}{\frac{\varphi_k}{\varphi_n}} \right) - \frac{1}{\theta_k} \right) \right]}}{\Lambda}$$

As the average markup, $\mathcal{M}$ depends on the speed of own-innovation relative to creative destruction θ, the productivity gaps and the distribution of productivity gaps.

4.6.1 Discussion of the stationary distribution

In equilibrium, high productivity firms expand into new markets at higher rates than low productivity firms as they charge higher markups upon market entry. Entry by new firms prevents high productivity firms from capturing all product lines. To see this note that in steady state $\dot{S} = 0$ such that eq. (10) turns into

$$z\left(S - p^h\right) = S(1-S)(x^h - x^l). \tag{14}$$

It is worthwhile to discuss eq. (14) since it provides intuition on the relationship between expansion rates and firm entry. If high productivity incumbents expand at higher rates than low productivity firms ($x^h > x^l$), for the share of high productivity incumbents to be constant along the BGP, S needs to be greater than p^h, the share of entrants with the high productivity type (given both types exist and there is positive firm entry). In other words, there must be "sufficient" low productivity firms entering such that the relatively higher expansion rate by existing high productivity incumbents is balanced and the share of lines operated by high productivity firms is constant. Eq. (14) highlights the role of firm entry. Without entry ($z = 0$) higher expansion rates by high productivity incumbents would result in those firms eventually overtaking all product lines. Firm entry is necessary to obtain a stationary distribution over productivity gaps: given $x^h > x^l$ and neither $S = 0$ nor $S = 1$, for eq. (14) to hold firm entry needs to be positive.

The relationship between differences in expansion rates across productivity types and firm entry in steady-state can most clearly be seen in the special case where all entrants are of the low productivity type ($p^h = 0$). In this case eq. (14) can be written as

$$Sx^h(1 - S) - (1 - S)x^l S = zS.$$

Entry by low productivity firms that replace high productivity firms (zS) makes up exactly for the lost market share of incumbent low productivity firms ($Sx^h(1 - S) - (1 - S)x^l S$) such that the aggregate share of high productivity firms remains constant.

4.7 Firm dynamics

In this section, I derive how firm markups, sales and survival evolve with firm age and I characterize the firm size distribution. The results of this section will again be used when taking the model to the data.

4.7.1 Markup dynamics

Firm markups or $TFPR$ are defined in eq. (5) as $\mu_f = \frac{py_f}{wl_f} = \left(\frac{1}{n_f}\sum_{i\in N_f}\mu_{if}^{-1}\right)^{-1}$. The firm markup is the harmonic mean of its product markups. In Appendix B.6 I show that for a high process efficiency firm the expected log markup conditional on firm age a_f is

$$E\left[\ln \mu_f | \text{firm age} = a_f, \varphi^h\right] = \underbrace{\ln \lambda \times \left(1 + I^h \times E[a_P^h|a_f]\right)}_{\text{Quality improvements}} + \underbrace{(1 - S) \times \ln\left(\frac{\varphi^h}{\varphi^l}\right)}_{\text{Productivity advantage}}, \tag{15}$$

where $E[a_P^h|a_f]$ denotes the average product age of a high process efficiency firm conditional on firm age, which is given by

$$\begin{aligned}
E[a_P^h|a_f] &= \frac{1}{x^h}\left(\frac{\frac{1}{\tau}\left(1 - e^{-\tau a_f}\right)}{\frac{1}{x^h+\tau}\left(1 - e^{-(x^h+\tau)a_f}\right)} - 1\right)\left(1 - \phi^h(a_f)\right) + a_f\phi^h(a_f) \\
\phi^h(a) &= e^{-x^h a}\frac{1}{\gamma^h(a)}\ln\left(\frac{1}{1 - \gamma^h(a)}\right) \\
\gamma^h(a) &= \frac{x^h\left(1 - e^{-(\tau - x^h)a}\right)}{\tau - x^h e^{-(\tau - x^h)a}},
\end{aligned}$$

Expected markups conditional on age consist of two terms. The first term in eq. (15) reflects how average quality gaps across the firm's products evolve with firm age. It captures the effects that as the firm ages it improves the quality of its continuing items, the firm loses product lines for which it has accumulated quality gaps and acquires new products with initially low quality gaps. This term is akin to the markup expression in Peters (2020). In Peters (2020), this term holds for all firms, whereas in my model this term is specific to the productivity type of the firm as own-innovation and expansion rates vary by firm type. Permanent differences in the process efficiency level across firms do not only affect expected markup growth (captured by the first term), but also introduce a level effect that is captured by the second term in eq. (15). The intuition behind the second term is that if a high process efficiency firm faces a low process efficiency second best producer in a given line, it can charge a φ^h/φ^l higher markup, which occurs in expectation in $1-S$ of the firm's product lines.

The expected markup conditional on firm age for a low process efficiency firm is

$$E\left[\ln \mu_f | \text{firm age} = a_f, \varphi^l\right] = \underbrace{\ln\lambda \times \left(1 + I^l \times E[a_P^l|a_f]\right)}_{\text{Quality improvements}} + \underbrace{S \times \ln\left(\frac{\varphi^l}{\varphi^h}\right)}_{\text{Productivity disadvantage}}. \tag{16}$$

The first term capturing the quality advantage is equivalent to the first term in eq. (15). $E[a_P^l|a_f]$ follows the same expression as $E[a_P^h|a_f]$ with h replaced by l. The second term in eq. (16) differs as low productivity firms have a process efficiency "dis"advantage in a share S of their product lines, where they face a high productivity second best producer. Since $\varphi^l < \varphi^h$, this term is negative.

4.7.2 Sales dynamics

Firms loose products according to the same stochastic process as in Klette and Kortum (2004) that, in their model, results in a skewed sales distribution, a decreasing variance of sales growth in size, a declining exit probability in age and size and firm size growth being independent of size. In my model, the rate at which firms add products is, however, heterogeneous across firms. High process efficiency firms add new products at a faster rate than low process efficiency firms, which in return affects firm sales and survival. Therefore, the properties related to firm size and survival in Klette and Kortum (2004) hold conditional on the productivity type of the firm.

Firm sales are proportional to the number of products a firm is producing. Since more productive firms add new products at a faster rate, sales growth is productivity type specific. Expected log sales growth for a firm with process efficiency $\varphi^j, j \in \{h, l\}$ between age zero and age a_f is $E\left[\ln nY | a_f, \varphi^j\right] - E\left[\ln nY | 0, \varphi^j\right]$, where n is the number of products a firm is producing. Firm sales growth stems from aggregate growth and from the firm gaining and loosing products as it ages

$$E\left[\ln nY | a_f, \varphi^j\right] - E\left[\ln nY | 0, \varphi^j\right] = \underbrace{g \times a_f}_{\text{Aggregate growth}} + \underbrace{E\left[\ln n | a_f, \varphi^j\right]}_{\text{Firm's product growth}}.$$

To derive $E\left[\ln n | a_f, \varphi^j\right]$ note that the probability of a high process efficiency firm producing n products at age a conditional on survival is $\left(1-\gamma^j(a)\right)\left(\gamma^j(a)\right)^{n-1}$, where $\gamma^j(a) = x^j \frac{1-e^{-(\tau - x^j)a}}{\tau - x^j e^{-(\tau-x^j)a}}$.

Therefore sales growth is given by

$$E\left[\ln nY|a_f,\varphi^j\right]-E\left[\ln nY|0,\varphi^j\right]=\underbrace{g\times a_f}_{\text{Aggregate growth}}+\underbrace{\left(1-\gamma^j(a_f)\right)\sum_{n=1}^{\infty}\ln n\times\left(\gamma^j(a_f)\right)^{n-1}}_{\text{Firm's product growth}}. \quad (17)$$

4.7.3 Employment dynamics and firm survival

Employment of the firm can be decomposed into the number of products that the firm produces and its markup as in Peters (2020). In my model product and markup dynamics depend on the process efficiency of the firm such that

$$E[\ln l_f|\text{firm age}=a_f,\varphi^j]=E\left[\ln\left(\frac{nY}{w\mu_f}\right)|a_f,\varphi^j\right]=\ln\left(\frac{Y}{w}\right)+E\left[\ln n|a_f,\varphi^j\right]-E\left[\ln\mu_f|a_f,\varphi^j\right],$$

where $\varphi^j\in\{\varphi^h,\varphi^l\}$. Employment growth then simply is

$$E[\ln l_f|a_f,\varphi^j]-E[\ln l_f|0,\varphi^j]=\underbrace{E\left[\ln n|a_f,\varphi^j\right]}_{\text{Firm's product growth}}-\underbrace{\left(E\left[\ln\mu_f|a_f,\varphi^j\right]-E\left[\ln\mu_f|0,\varphi^j\right]\right)}_{\text{Firm's markup growth}}. \quad (18)$$

$E\left[\ln\mu_f|a_f,\varphi^j\right]-E\left[\ln\mu_f|0,\varphi^j\right]$ and $E\left[\ln n|a_f,\varphi^j\right]$ are derived in eq. (15), (16) and (17).

Firm size dynamics determine firm survival since firms that loose their last product become inactive. Firms that grow faster in size have a higher probability of survival. Hence, firm survival is productivity type specific. The share of high and low process efficiency firms surviving until age a is given by

$$s^h(a_f)=1-\tau\frac{1-e^{-(\tau-x^h)a_f}}{\tau-x^he^{-(\tau-x^h)a_f}} \quad (19)$$

$$s^l(a_f)=1-\tau\frac{1-e^{-(\tau-x^l)a_f}}{\tau-x^le^{-(\tau-x^l)a_f}}. \quad (20)$$

4.7.4 Firm size distribution

The mass of high and low process efficiency firms with $n\geq 2$ products follows the differential equations

$$\begin{aligned}\dot{M}_t^h(n)&=(n-1)x^hM_t^h(n-1)+(n+1)\tau M_t^h(n+1)-n(x^h+\tau)M_t^h(n)\\ \dot{M}_t^l(n)&=(n-1)x^lM_t^l(n-1)+(n+1)\tau M_t^l(n+1)-n(x^l+\tau)M_t^l(n),\end{aligned} \quad (21)$$

whereas the mass of firms with one product evolves according to

$$\begin{aligned}\dot{M}_t^h(1)&=zp^h+2\tau M_t^h(2)-(x^h+\tau)M_t^h(1)\\ \dot{M}_t^l(1)&=z(1-p^h)+2\tau M_t^l(2)-(x^l+\tau)M_t^l(1).\end{aligned} \quad (22)$$

The mass of firms with n products increases through firms with $n-1$ products expanding to size n at rate x^h or x^l per product or through firms with $n+1$ products losing a product at the rate of aggregate creative destruction. The mass of firms with n products decreases through firms with n

products either losing or gaining a product through expansion or creative destruction. The mass of firms with one product additionally increases through firm entry. The change in the total mass of firms at any point in time is $z - \tau(M_t^h(1) + M_t^l(1))$. To find the stationary firm size distribution I set the time derivatives in equations (21) and (22) equal to zero. The steady-state mass of high and low process efficiency firms with n products is

$$M^h(n) = \frac{(x^h)^{n-1} z p^h}{n\tau^n} = \frac{zp^h}{x^h}\frac{1}{n}\left(\frac{x^h}{\tau}\right)^n$$
$$M^l(n) = \frac{(x^l)^{n-1} z(1-p^h)}{n\tau^n} = \frac{z(1-p^h)}{x^l}\frac{1}{n}\left(\frac{x^l}{\tau}\right)^n.$$

From this one obtains (all in steady-state) the total mass of firms with n products

$$M(n) = M^h(n) + M^l(n) = \frac{(x^h)^{n-1} z p^h + (x^l)^{n-1} z(1-p^h)}{n\tau^n}, \tag{23}$$

the mass of firms of each productivity type

$$M^h = \sum_{n=1}^{\infty} M^h(n) = \frac{zp^h}{x^h}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{x^h}{\tau}\right)^n = \frac{zp^h}{x^h}\ln\left(\frac{\tau}{\tau - x^h}\right)$$
$$M^l = \sum_{n=1}^{\infty} M^l(n) = \frac{z(1-p^h)}{x^l}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{x^l}{\tau}\right)^n = \frac{z(1-p^h)}{x^l}\ln\left(\frac{\tau}{\tau - x^l}\right)$$

and the total mass of firms

$$M = M^h + M^l.$$

As in Lentz and Mortensen (2008), the share of firms with n products for each firm type, $M^h(n)/M^h$ and $M^l(n)/M^l$, follows the pdf of a logarithmic distribution with parameter x^h/τ and x^l/τ. The firm size distribution is highly skewed to the right.

Since there is a continuum of mass one of products and each product is mapped to one firm $\sum_{i=1}^{\infty} M(n) \times n = 1$. Further, the mass of high process efficiency firms producing n products is related to the previously defined share of product lines operated by high process efficiency firms, S, as follows

$$S = \sum_{n=1}^{\infty} M^h(n) \times n = \frac{zp^h}{\tau - x^h}.$$

From the firm size distribution I obtain the share of high process efficiency firms

$$S_{M^h} = \frac{M^h}{M},$$

the entry rate

$$\text{Entry rate} = \frac{z}{M} \tag{24}$$

and the standard deviation of log firm sales

$$\sigma(\ln sales) \equiv \sigma(\ln pyn) = \sqrt{\sum_{n=1}^{\infty}\frac{M(n)}{M}(\ln n)^2 - \left(\sum_{n=1}^{\infty}\frac{M(n)}{M}\ln n\right)^2}. \tag{25}$$

4.8 Testing model predictions

4.8.1 Quality, price and revenue productivity growth

In the model, product innovations occur along the quality dimension, i.e., every product innovation leads to a λ improvement in product quality. A quality improvement through incumbent own-innovation increases the product price and markup according to eqs. (2) and (3). Linking micro data on firms' price setting with balance sheet data, I provide suggestive evidence that quality improvements are positively related to price growth at the product level and that firms with higher price growth display faster revenue productivity growth. Taken together the results suggest that innovations (quality improvements) are associated with faster revenue productivity (or markup) growth.

To study the price setting of firms I exploit the micro data underlying the Swedish Producer Price Index (PPI) that contains reported item prices and information on quality-adjusted prices. If firms report to Statistics Sweden (SCB) that the nature of their good has changed, e.g., due to changes in quality, SCB assesses the change in quality and computes a quality-adjusted price for the new item. If the new item is an updated version of the old item, SCB keeps the item ID unchanged. I identify a quality change as an instance where SCB updates the quality-adjusted price, while the item ID remains unchanged. Around the time of the quality improvement, the nominal price is expected to increase according to the theory.

Table 2: Change in nominal price after quality improvement

	$\ln \frac{p_t}{p_{t-1}}$
$\mathbb{1}\{Quality\ adjustment\}_t$	0.1833
	(5.193)
Constant	0.0179
	(73.802)
Industry × year FE	✓
N	971,467

Notes: t statistics in parentheses. Clustered standard errors at the industry × year level. $\frac{p_t}{p_{t-1}}$ denotes annualized month-to-month price growth. Number of quality adjustments for continued items in data: 6,676. Regression restrited to items with an annualized monthly price growth rate of $|\ln p_{ikt,month} - \ln p_{ikt,month-1}| < 0.5 \times 12$, which includes 99.5% of the observations.

To test whether changes in item quality are reflected in nominal price changes I regress (annualized) monthly log price growth rates at the item level, i, on a dummy indicating whether a quality adjustment occurred in that month controlling for industry × year fixed effects, δ_{kt}:

$$\ln \frac{p_{ikt,month}}{p_{ikt,month-1}} = \beta_0 + \beta_1 \mathbb{1}\{Quality\ adjustment\}_{ikt,month} + \delta_{kt} + u_{ikt,month}.$$

The results are shown in Table 2. The coefficient on the quality adjustment dummy is positive and statistically significant. Items with changes in quality display an annualized 0.18 log point higher price increase relative to non-adjusted items. In line with the theory, quality adjustments are associated with price increases.

To show that firms with higher price growth display faster $TFPR$ growth I regress log $TFPR$ changes at the firm-level on price changes controlling for industry × year fixed effects.

$$\ln \frac{TFPR_{fkt}}{TFPR_{fkt-1}} = \beta_0 + \beta_1 \ln \frac{p_{fkt}}{p_{fkt-1}} + \delta_{kt} + u_{fkt}$$

Since prices are observed at the item level at monthly frequency in the PPI micro data I compute annual firm price growth rates as sales weighted December to December item price growth rates of the firm. This allows me to relate price changes to $TFPR$ changes that are computed at the firm-level at annual frequency in the balance sheet data.

Table 3: Revenue productivity and output price growth

	$\ln \frac{TFPR_t}{TFPR_{t-1}}$
$\ln \frac{p_t}{p_{t-1}}$	0.2395
	(4.737)
Industry × year FE	✓
N	15,673

Notes: t statistics in parentheses. Clustered standard errors at the industry × year level. $\ln \frac{TFPR_{fkt}}{TFPR_{fkt-1}}$ denotes a firm's log TFPR growth with respect to the previous year. $\frac{p_t}{p_{t-1}}$ denotes a firm's December to December price growth averaged over the products of the firm (sales weighted).

Table 3 shows the results. The coefficient on price changes is positive and significant. Larger price changes are associated with higher $TFPR$ growth. To be precise, firms with one log point higher price growth display on average 0.24 log points larger $TFPR$ growth. One comment is in order. The PPI micro data covers a sample and not the universe of item prices. The computed firm price growth rate therefore only reflects items that are included in the PPI and potentially does not include all items of the firm. If the computed firm price growth rate is equal to the true firm price growth rate plus a measurement error and true price growth rates are uncorrelated with the measurement error, the regression coefficient is biased towards zero and should be viewed as a lower bound. Another reason why the estimate should be interpreted as a lower bound is that unobserved productivity changes of the firm are arguably negatively related to prices changes and positively to $TFPR$ changes. Hence, price growth that originates from quality improvements potentially has larger effects on $TFPR$ growth than the one shown in Table 3.

The two regressions suggest that quality improvements lead to faster price growth and that firms with higher price growth increase their revenue productivity faster. The two facts link innovation with revenue productivity growth through price growth.

4.8.2 Determinants of revenue productivity

In the model, revenue productivity increases as the firm ages through the process of own-innovation. This suggests that revenue productivity is positively related to firm age. Simultaneously, more productive firms (with higher revenue productivity) grow larger in size. This suggests that revenue productivity is positively related to firm size across firms. In Appendix C.1 I run a horse-race regression of revenue productivity on firm age and firm size. Consistent with the model, the regression identifies both firm age and firm size as independent drivers of revenue productivity across firms. I further show that the revenue productivity-size relation within firms is smaller than across firms and that the revenue productivity-size relationship when size is measured by employment is smaller than when size is measured by sales. Both findings are consistent with the theory.

5 Model application

In this section I study the drivers behind the trends in the Swedish economy. I estimate the model using Swedish administrative data in a first step.

5.1 Estimation

There are in total nine parameters in the model. The own-innovation efficiency for high and low productivity firms ψ_I^h, ψ_I^ℓ, the expansion efficiency ψ_x, the innovation cost curvature ζ, the entry efficiency ψ_z, the step size of innovation λ, the process efficiency differential $\frac{\varphi^h}{\varphi^\ell}$, the probability of being assigned the high process efficiency type conditional on entry p^h, and the discount factor ρ. I follow Acemoglu, Akcigit, Alp, Bloom and Kerr (2018) and Peters (2020) that set ζ equal to two based on evidence from the microeconometric innovation literature (Blundell, Griffith and Windmeijer, 2002; Hall and Ziedonis, 2001). The discount factor ρ is set to 0.05. In the estimation, I target firm life cycle growth that, in the model, is productivity type specific. Since firm productivity is unobserved in the data, I follow a model based approach to classify firms into productivity types. The model predicts that firms with a high process efficiency expand faster in size. I identify high productivity firms in the data as firms that are large on average conditional on firm age. In a first step I regress firm sales on firm age, industry fixed effects and firm fixed effects. I then split firms into two groups according to their firm fixed effect and interpret firms with relatively high firm fixed effects as high process efficiency firms. Knowing the productivity type of a firm, I then compute p^h from the data as the share of firms at age zero that are of the high process efficiency type. I set the threshold for the firm fixed effect in the productivity classification such that 65% of the entrants are of the high type.[12] This leaves in total six parameters to estimate $(\psi_I^h, \psi_I^\ell, \psi_x, \psi_z, \lambda, \frac{\varphi^h}{\varphi^\ell})$.

I target six moments both at the firm-level and at the aggregate-level. At the firm-level, I match the process efficiency specific life cycle growth of $TFPR$ and employment (four moments in total).[13] I target employment growth between age zero and age 15, whereas for $TFPR$ I target growth between ages zero and ten. $TFPR$ growth displays strong concavity. To match $TFPR$ growth at early years of the firm I take age ten as the target. I further match the $TFPR$ difference between high and low process efficiency firms at age two to discipline the process efficiency differential.[14]

[12] In section 3 and the appendix, I show that the trends in the Swedish economy are not driven by a small share of firms, but the trends happen along the entire firm distribution. I therefore choose a more equal split of firms into productivity types than, e.g., focusing on the top 10% of firms. I check robustness with respect to the threshold.

[13] Both employment and sales growth are natural candidates to target in the estimation. Using employment has the following advantages: sales levels of the firm are already used to classify firms into productivity types. Even though I target growth rates in the estimation, I expect less implications of the productivity classification for employment than for sales growth. Sales might be further related to value added that shows up in the numerator of $TFPR$. $TFPR$ growth is already targeted in the calibration. The employment measure is taken from a separate data source (RAMS).

[14] The reason I take age two to measure the gap is that this is the age when I observe most firms in the data.

The $TFPR$ and employment life cycles as well as the difference in $TFPR$ between high and low process efficiency firms are characterized analytically in eq. (15), (16) and (18). I obtain $TFPR$ and employment life cycle growth from the data by regressing log $TFPR$ and log employment (net of industry $\times$ year fixed effects) on firm age dummies (age zero left out) and a constant. For high process efficiency firms, $TFPR$ between age zero and ten grows by 0.114 log points and employment between age zero and 15 grows by 0.524 log points. For low process efficiency firms I observe $TFPR$ growth of 0.098 log points and employment growth of 0.123 log points. The level difference of $TFPR$ between high and low process efficiency firms at age two is 0.046 log points. At the aggregate level I target a TFP growth rate of 2.3%, which is taken from the Multifactor productivity (MFP) series by Statistics Sweden (SCB) measured in labor augmenting terms for the period 1996–2006. This growth rate is characterized analytically by eq. (13). Since all moments are derived analytically as functions of the endogenous model outcomes, there is no need to simulate firm life cycles for the estimation. I estimate the model using the generalized method of moments (GMM).[15]

The moment conditions and calibrated parameters are shown in Table 4. High process efficiency firms are roughly 40% more efficient in own-innovation than low process efficiency firms: $\psi_I^h/\psi_I^l = 1.39$, which is estimated from their steeper $TFPR$ growth profile. Firms with the high process efficiency type are roughly 4% more productive than low process efficiency firms: $\varphi^h/\varphi^l = 1.039$. Each product innovation improves the quality of an item by roughly 6.3%: $\lambda = 1.063$.

5.2 Model fit of the Swedish economy

Table 4 shows that all targeted moments are precisely hit by the model. How does the model compare along dimensions of the Swedish economy that are not directly targeted? Figure 4 shows the life cycles of $TFPR$, employment and sales growth both for the data and the model. Only $TFPR$ growth with respect to age ten and employment growth with respect to age 15 is targeted, the remaining years are not. Sales life cycle growth is not targeted at all. The displayed life cycle growth profiles are obtained from the data as explained in the previous section. In panel (a), I relate $TFPR$ growth from the data to the $TFPR$ level in the model by normalizing the growth profile to the model implied $TFPR$ level for high and low process efficiency firms at age zero.

Overall the model fits the growth profiles for $TFPR$, employment and sales well. For $TFPR$, the model overshoots later years of the firm slightly, particularly for low process efficiency firms. However, at age 20 only 16% of the low and 27% of the high process efficiency firms are still active such that discrepancies between data and model for later years affect only a small share of firms. The employment fit is very good for both productivity types. Recall that the classification of the process efficiency type is based on sales and the measure of employment is taken from an independent data source (RAMS). Almost all data points align with the model implied employment growth. Even though not targeted directly, the model fits the sales life cycle particularly for the later years. I further compare the size distribution in the model and in the data, panel (d). The size distribution is displayed in form of the Lorenz curve, capturing the share of value added that is attributed to

Age in my data is imputed based on the first employee of the firm, which is taken from a data set (RAMS) that is independent of the balance sheet data used to compute $TFPR$. Some firms don't report balance sheets the year they hire their first employee and hence show up in the balance sheet data after age zero.

[15] All targets receive equal weight. I use Julia's non-linear solver NLsolve with the trust-region method to minimize the distance between model and data targets.

Table 4: Targeted moments and estimated parameters

		Data	Model	Source
Targets				
	$TFPR$ growth high prod.	0.114	0.114	Own calc.
	$TFPR$ growth low prod.	0.098	0.098	Own calc.
	Employment growth high prod.	0.524	0.524	Own calc.
	Employment growth low prod.	0.123	0.123	Own calc.
	$TFPR$ level difference	0.046	0.046	Own calc.
	Agg. growth rate g	2.3%	2.3%	SCB
Parameters				
ψ_I^h	*Inno. efficiency high prod.*		1.504	Estimated
ψ_I^l	*Inno. efficiency low prod.*		1.081	Estimated
ψ_x	*Expansion efficiency*		0.229	Estimated
ψ_z	*Entry efficiency*		1.336	Estimated
λ	*Step size of inno.*		1.063	Estimated
φ^h/φ^l	*High/low productivity gap*		1.039	Estimated
ρ	*Discount rate*		0.05	Set exog.
ζ	*Inno. cost curvature*		2	Set exog.
p^h	*Prob. high prod. type (entry)*		0.653	Set exog.

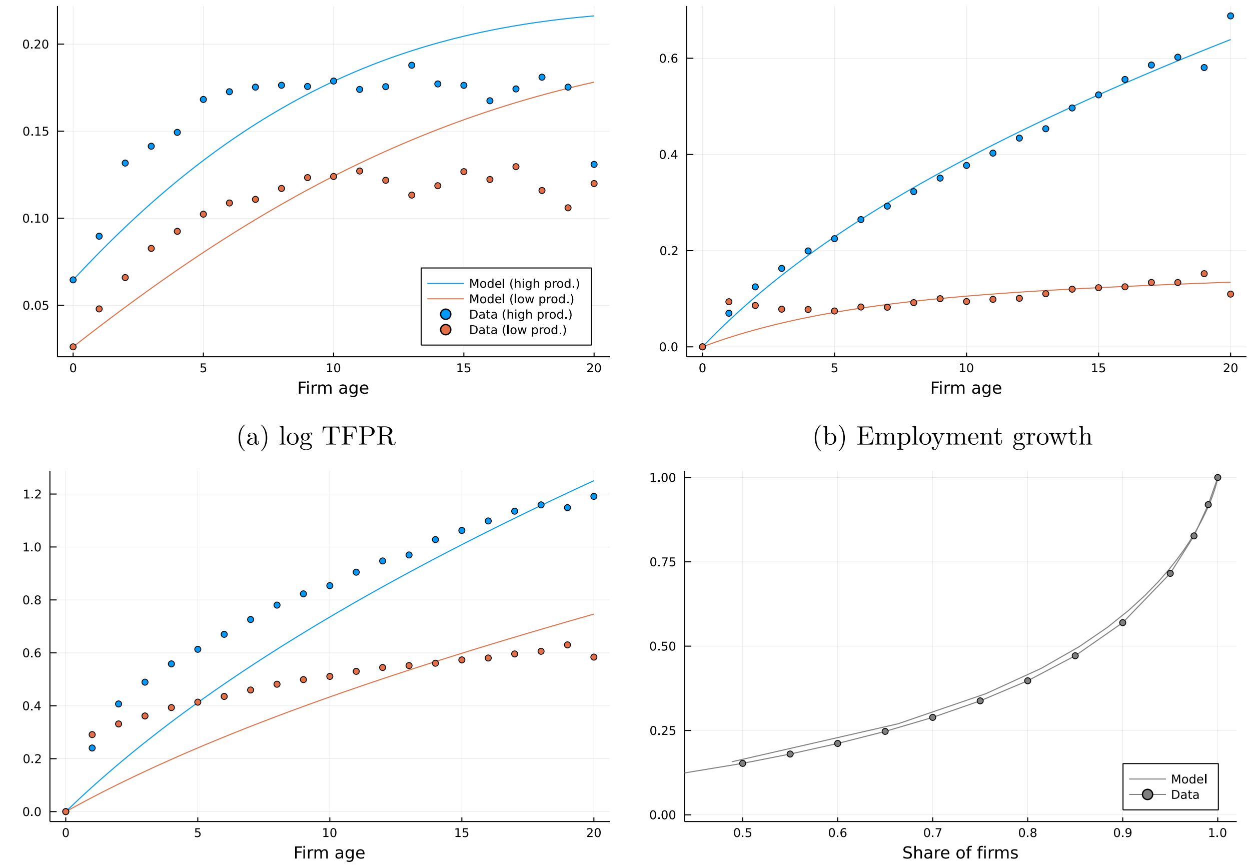
Notes: firm life cycle growth and p^h obtained from Swedish administrative data. Life cycle growth measured as log difference between age zero and ten ($TFPR$) or age zero and 15 (employment). The $TFPR$ difference between high and low productivity firms is measured at age two. The agg. growth target is obtained from Statistics Sweden (SCB), MFP series (1996–2006) in labor augmenting terms.

the smallest $x\%$ of firms. Both model and data show very similar levels of concentration for output well below the imaginary 45-degree line.

$TFPR$ growth of high productivity firms is only slightly higher than $TFPR$ growth of low productivity firms both in the data and in the estimated model (11.4% vs. 9.8% over the first ten years). Note that this does not imply that both firm types own-innovate at a similar speed. More productive firms own-innovate and expand into new markets at faster rates than less productive firms. The fact that more productive firms add new product lines with low markups to their product portfolio at a faster rate deflates their $TFPR$ growth. Their higher expansion rate therefore partly offsets the effect of faster own-innovation on firm $TFPR$ growth. Their stark difference in expansion rates can be seen in the heterogeneous life cycle profiles for sales in Figure 4.

The model also compares well along other non-targeted dimensions. The model implies an equilibrium entry rate, characterized by eq. (24), of 0.059. Using the same administrative records as I do in this paper, Engbom (2020) reports an entry rate for the Swedish economy of around 0.065 in the late 1980s followed by a drop to 0.05 to 0.055 during 2010–2015. In terms of markups, the model generates a sales-weighted markup average of 1.19 across product lines. The cost-weighted markup

Figure 4: Model fit of firm growth and size distribution (untargeted)



(a) log TFPR

(b) Employment growth

(c) Sales growth

(d) Share of value added

Notes: figure shows the fit of life cycle growth for high and low process efficiency firms separately as well as the fit of the firm size distribution. $TFPR$, employment and sales growth obtained from the data by regressing its log (net of industry $\times$ year fixed effects) on age dummies (age zero left out) and a constant. Coefficients indicate the log difference to age zero. $TFPR$ growth is normalized by the model implied log $TFPR$ at age zero for each productivity type in panel (a). In the above panels only $TFPR$ growth at age ten and employment growth at age 15 are targeted in the calibration.

is 1.16. Sandström (2020) documents the time series of sales-weighted average markups in Sweden using the same underlying data source as I do for firms with more than ten employees. Average markups range within 1.13 to 1.23 during 1998–2016. De Loecker and Eeckhout (2018) report a sales-weighted markup of 1.31 for large and mainly publicly traded firms in Sweden in 2016. The distribution of sales across productivity types is also similar in the model and in the data. The sales share of high process efficiency firms is 0.91 in the model (S) vs. 0.92 in the data measured as the sales share of high process efficiency firms in the cross section of firms averaged across all years 1997–2017.

5.3 Explaining the recent economic trends

In this section I apply the model to explain the 19.5% increase in the standard deviation of log $TFPR$, the 5% increase in the standard deviation of log sales shares and the 25% increase in

$TFPR$ life cycle growth at age four. Starting from the calibrated balanced growth path (BGP) I re-estimate model parameters to replicate the trends along a new BGP. I focus on changes in the cost of own-innovation for the high productivity firms ψ_I^h, the cost of creative destruction ψ_x and the entry cost ψ_z. Changes in the innovation costs affect the firms' optimal innovation efforts directly through the optimality conditions and are therefore particularly well-suited to match the changes in $TFPR$ dispersion, sales concentration and $TFPR$ life cycle growth quantitatively. A higher own-innovation rate increases $TFPR$ life cycle growth and $TFPR$ dispersion ceteris paribus. A faster rate of firm expansion leads to more sales concentration. The rate of creative destruction limits the extent of life cycle growth and cross sectional dispersion in $TFPR$ and sales. ψ_I^h, ψ_x and ψ_z therefore match changes in $TFPR$ life cycle growth as well as dispersion in $TFPR$ and sales quite flexibly. Substituting one of the parameters results in the estimation falling short in explaining the three trends quantitatively as explained later on.

The 25% increase in $TFPR$ life cycle growth at age four in Figure 3b is obtained from the pooled sample of firms, i.e., firms of both productivity types. I therefore target a 25% increase in $TFPR$ life cycle growth for the pooled sample of firms in the model, weighting productivity type specific $TFPR$ growth obtained from eqs. (15) and (16) with the type specific entry probability (p^h and $1-p^h$) and survival probability (eqs. (19) and (20) all evaluated at age four). The standard deviation of log sales is characterized by eq. (25). To compute the standard deviation of $TFPR$ across firms, I simulate the model.[16] I re-estimate $(\psi_I^h, \psi_x, \psi_z)$ using the simulated method of moments (SMM) matching the industry and firm level trends along a new BGP.[17]

Table 5: New balanced growth path

	Parameters			*Moments*		
	ψ_z	ψ_I^h	ψ_x	$\sigma(\ln TFPR)$	$\sigma(\ln sales)$	$TFPR$ life cycle
Data				+19.5%	+5.0%	+25.0%
Model	-11.1%	+30.8%	-12.8%	+19.7%	+5.0%	+25.0%

Notes: the table shows the estimated parameter changes. ψ_I^h denotes the own-innovation efficiency (high type), ψ_x the creative destruction efficiency and ψ_z the entry efficiency.

Table 5 shows the estimated changes for the three parameters. The estimation asks for an 11% increase in the entry costs, a 31% decrease in the cost of own-innovation for the high productivity type and a 13% increase in the cost of creative destruction. The three parameter changes are able to replicate the change in $TFPR$ life cycle growth, $TFPR$ dispersion and sales concentration vis-à-vis the initial BGP quantitatively. The changes in own-innovation and creative destruction costs capture the general notion that for more productive firms to improve their own-products in markets where they have already established themselves as the market leader has become easier. At the same time, for firms (incumbents and entrants) to expand into new product markets where competing firms have already established themselves as the market leader has become harder.

[16] I simulate the economy for 200 years, the last period is used to compute the standard deviation of log $TFPR$ across firms. Since time is continuous in the model I discretize time into 50 intervals within a year and simulate an economy consisting of 15.000 products.

[17] All three targets receive equal weight. I use Julia's non-linear solver NLsolve with the trust-region method to minimize the distance between the three model and data targets.

The changes in the innovation costs replicate the rise in misallocation, market concentration and acceleration in firm markup growth as follows. The decrease in the own-innovation costs for the more productive firms incentivizes those firms to own-innovate faster. A rise in the own-innovation efforts accelerates markup accumulation of those firms, i.e., markup life cycle growth increases. The acceleration in markup life cycle growth is associated with an increase in cross-sectional markup dispersion generating a rise in misallocation. At the same time, as more productive firms accumulate markups faster they have a higher incentive to enter new product markets. Together with a fall in firm entry that is due to an increase in the creative destruction costs for entrants this increases market concentration. Somewhat surprising is the increase in the cost of firm expansion for incumbent firms ψ_x. To generate an increase in sales concentration one might expect a decrease in the cost of firm expansion. The reason why the estimation asks for an increase in the cost of firm expansion is that the increase in the entry cost more than explains the rise in sales concentration as illustrated in Table 6.

Table 6: Decomposing the rise in sales concentration

	Baseline	$(\psi_z \downarrow, \psi_I^h \uparrow, \psi_x \downarrow)$	$\psi_z \downarrow$	$\psi_I^h \uparrow$	$\psi_x^* \downarrow$
Sales concentration $\sigma(\ln sales)$	0.836	+5.0%	+44.8%	-15.3%	-45.5%

Notes: the table shows the effect of each parameter change on sales concentration in isolation. The Baseline column shows levels in the initial BGP, remaining columns show percentage changes. $\psi_z \downarrow$ denotes the 11.1% decline in the entry efficiency, $\psi_I^h \uparrow$ the 30.8% increase in the own-innovation efficiency (high type) and $\psi_x \downarrow$ the 12.8% decrease in the expansion efficiency. $\psi_x^* \downarrow$ shows the linearized response to the 12.8% decrease in ψ_x as $\psi_x \downarrow$ does not generate a BGP solution in isolation.

Table 6 shows the change in the standard deviation of log sales in response to the parameter changes (both jointly and in isolation). Starting from the new BGP I obtain the effect of each parameter change in isolation by turning the parameter change off (and on). The increase in the entry cost leads to a 44.8% increase in the standard deviation of log sales. The decline in creative destruction that results from the increase in entry costs raises firms' expansion efforts by more than what is required to match the targeted 5% increase in sales concentration. As a result, the estimation asks for an increase in the firm expansion costs. This highlights the importance of replicating the empirical trends jointly as a decrease in the expansion costs generates a rise in concentration in isolation.

In the estimation, I have focused on changes in own-innovation, creative destruction and entry costs. To test for alternative drivers of the empirical trends I re-run the SMM estimation for all possible parameter triples in the model. It turns out that other parameter combinations, despite replicating the trends qualitatively, fall short in explaining them quantitatively. Alternative parameter combinations that get close in replicating the trends quantitatively consistently feature a decrease in the own-innovation costs for the high type, an increase in entry costs or an increase in the creative destruction costs, which is the reason why those three parameters in combination fit the trends particularly well. Firm type heterogeneity also plays an important role. Own-innovation costs for the high productivity firms decrease, while the costs for the low productivity firms remain unchanged. Decreasing the own-innovation costs for the low productivity firms by the same amount as for the high type firms (alongside the other parameter changes) results in a decrease in $\sigma(\ln sales)$ by 9%. The reason that sales concentration decreases is that lowering the cost of own-innovation for the low productivity type incentivizes those firms to expand in size, thereby stealing sales shares from the high productivity firms.

5.4 Implications for the long-run economy

In response to the parameter changes, the model makes predictions about the entry rate, markups, misallocation and the aggregate growth rate in the long-run, see Table 7. The model predicts a decline in the entry rate, which is consistent with Engbom (2020) who, using the same data source as I do in this paper, documents a decline in the firm entry rate in Sweden that goes back to the 1980s. The model further predicts an increase in the aggregate cost weighted markup from roughly 1.16 to 1.2. As explained in the previous section, the changes in innovation costs incentivize more productive firms to accumulate markups faster and to expand into new product markets. In other words, the firms that accumulate markups faster are the ones that increase their sales shares. This is in line with Kehrig and Vincent (2021) who show that the fall in the aggregate labor income share in the US is generated by firms whose labor share fell as they grew in size. In my model, the aggregate labor income share is the inverse of the aggregate (cost weighted) markup.

Static efficiency in the model is affected through changes in both markup dispersion and aggregate productivity. The output loss that arises from markup dispersion increases from roughly 1.2% to 1.9% ($\mathcal{M}$ falls by roughly 0.7%). The fall in the own-innovation costs is responsible for the rise in markup dispersion: as firms accumulate markups faster, misallocation through markup dispersion increases. At the same time, more productive firms take over product lines from less productive firms resulting in an increase in aggregate productivity by 0.06%. The increase in aggregate productivity is small compared to the increase in markup dispersion such that overall static efficiency falls.

Table 7: Implications for the long-run economy

	Initial BGP	$(\psi_z \downarrow, \psi_I^h \uparrow, \psi_x \downarrow)$
Aggregates		
Entry rate	0.059	-0.23pp
Agg. markup $E^{cost}[\mu] - 1$	0.159	+4.23pp
Static efficiency		
Markup dispersion $\mathcal{M}$	0.988	-0.68%
Agg. productivity Φ		+0.06%
Economic growth		
Growth rate g	0.023	+0.48pp

Notes: the table shows the effect of the estimated parameter changes on the long-run economy. First column shows levels in the initial balanced growth path. $(\psi_z \downarrow, \psi_I^h \uparrow, \psi_x \downarrow)$ shows changes in percentage points (pp) or in percent after the rise in entry costs, fall in own-innovation costs and rise in creative destruction costs.

Interestingly, the model predicts an increase in the aggregate growth rate from 2.3% to roughly 2.8%. Why is the aggregate growth rate increasing? To understand the increase in economic growth, I run counterfactuals changing each innovation cost in isolation. Table 8, last row, shows the change in the aggregate growth rate Δg in response to each parameter change. The change in own-innovation costs for the more productive firms accounts for the largest fraction of the change in growth. The more productive firms increase their own-innovation efforts in response to the changes in the own-innovation costs, thereby generating economic growth. The rise in own-innovation causes both misallocation through markup dispersion and economic growth. This creates two opposing effects. Rising misallocation lowers the level of aggregate output permanently, which works in the opposite direction as the increase in the growth rate of aggregate output. The rise in the creative destruction costs for entrants, $\psi_z \downarrow$, also contributes a positive growth effect (second column, Table 8) by incentivizing incumbent firms to innovate faster. Most of the increase in the aggregate growth rate is, however, due to the decrease in own-innovation costs.

More productive firms (with higher innovation rates) increase their innovation efforts and expand into new product markets. I quantify how much of the increase in the aggregate growth rate is coming from changes in economic growth generated in the average product line and reallocation of product lines towards more innovative firms. The aggregate growth rate in eq. (13) can be written as

$$g = Sg^h + (1 - S)g^l,$$

where $g^h \equiv (I^h + x^h + z)\ln(\lambda)$ and $g^l \equiv (I^l + x^l + z)\ln(\lambda)$ denote the growth rate per product line, where a high or low productivity type firms is the incumbent. Using a standard shift-share decomposition I decompose changes in g as follows

$$\Delta g = g_{new} - g_{old} = \underbrace{S_{old}\Delta g^h + (1 - S_{old})\Delta g^\ell}_{\Delta\text{Within}} + \underbrace{g^h_{old}\Delta S - g^\ell_{old}\Delta S}_{\Delta\text{Between}} + \underbrace{\Delta g^h\Delta S - \Delta g^\ell\Delta S}_{\Delta\text{Cross term}}, \tag{26}$$

where *old* and *new* index steady-state variables before and after the parameter change. Changes in the aggregate growth rate are due to changes in innovation efforts per line holding the composition of product lines across firm types constant (ΔWithin), due to changes in the composition of product lines across firms of different type holding innovation efforts per line constant (ΔBetween) and due to changes in both innovation efforts and the composition of lines (ΔCross-term). The ΔCross-term turns out to be small such that I group ΔBetween and ΔCross-term into a common ΔReallocation term.

Table 8 shows the decomposition of changes in the aggregate growth rate between the initial and new BGP into the contributions from ΔWithin and ΔReallocation. Of the total change in aggregate growth (47.6bp) the majority comes from the ΔWithin effect (45.6bp). The ΔReallocation effect is relatively small. The reason is that the decline in own-innovation costs for the high type leads to a strong ΔWithin response: as own-innovation becomes cheaper own-innovation efforts in the average product line increase. This does not, however, imply that ΔReallocation effects are generally small compared to ΔWithin effects in the estimated model. Column 2 shows the decomposition for the entry cost increase in isolation. The ΔReallocation effect is ten times the size of the ΔWithin effect: the increase in entry costs affects growth mainly through reallocation as more innovate firms capture more product lines. The increase in creative destruction costs (last column) also

Table 8: Decomposing the change in the agg. growth rate

	$(\psi_z \downarrow, \psi_I^h \uparrow, \psi_x \downarrow)$	$\psi_z \downarrow$	$\psi_I^h \uparrow$	$\psi_x^* \downarrow$
ΔWithin	45.6bp	1.6bp	44.6bp	4.1bp
ΔReallocation	2.0bp	11.6bp	-3.4bp	-11.2bp
Δg	47.6bp	13.3bp	41.2bp	-7.1bp

Notes: the table shows the effect of each estimated parameter change on the agg. growth rate and decomposes changes in growth into ΔWithin and ΔReallocation effects. ΔWithin and ΔReallocation follow the decomposition in eq. (26). $\psi_z \downarrow$ denotes the 11.1% decline in the entry efficiency, $\psi_I^h \uparrow$ the 30.8% increase in the own-innovation efficiency (high type) and $\psi_x \downarrow$ the 12.8% decrease in the expansion efficiency. $\psi_x^* \downarrow$ shows the linearized response to the 12.8% decrease in ψ_x as $\psi_x \downarrow$ does not generate a BGP solution in isolation.

leads to ΔReallocation effects that dominate the ΔWithin effect. In this case, the increase in creative destruction costs reallocates market shares to low productivity incumbents with lower innovation rates such that the ΔReallocation effect is negative. The estimated change in the creative destruction cost and the entry cost are small compared to the change in the own-innovation cost, however, such that most of the total change in aggregate growth is coming from the ΔWithin response.

The changes in innovation costs imply a fall in static efficiency and an increase in the aggregate growth rate of the economy. To trade off the changes in efficiency and growth, I compare the change in lifetime utility across the old and the new BGP. In particular, I compare lifetime utility of BGPs, starting from the same level of average quality $\mathcal{Q}$ (see Appendix E for a detailed description). Lifetime utility of a consumption path where C_t grows at rate g is

$$\mathcal{U}\left(\{C_t\}_{t=0}^{\infty}\right) = \int_0^{\infty} e^{-\rho t} \ln C_t dt = \frac{1}{\rho} \ln C_0 + \frac{g}{\rho^2} = \mathcal{U}\left(C_0, g\right).$$

Lifetime utility along the BGP depends on three variables: the fixed parameter ρ that captures the household's discount rate, detrended consumption C_0 and the growth rate of the economy g. In the new BGP, detrended consumption is 2.6% lower than in the initial BGP. Both a fall in static efficiency and production labor reduce detrended consumption. The upside is that the growth rate of the economy is 0.48pp higher in the new BGP. I measure the change in lifetime utility in permanent consumption-equivalent terms, ξ, as follows

$$\mathcal{U}\left((1+\xi)C_0^{old}, g^{old}\right) = \frac{\ln(1+\xi)}{\rho} + \mathcal{U}\left(C_0^{old}, g^{old}\right) = \mathcal{U}\left(C_0^{new}, g^{new}\right). \tag{27}$$

ξ measures the permanent increase in consumption in the *old* BGP that equates lifetime utility of the *old* and *new* BGP. The change in innovation costs is associated with a 7.1% permanent increase in consumption, $\xi = 0.071$. Note that this comparison does not take into account the transition from the initial to the new BGP.

5.5 Policy counterfactuals

The change in innovation costs are associated with a long-run increase in misallocation. To illustrate to what extent a creative destruction subsidy can counteract the rise in misallocation, I run policy counterfactuals. In particular, starting from the initial BGP, I introduce a government that taxes firm profits and subsidizes the firms' creative destruction expenditures, while running a balanced budget.[18] A subsidy on creative destruction costs reduces misallocation as follows. The cost reduction incentivizes firms to engage in creative destruction. An increase in the aggregate creative destruction rate lowers the average life expectancy of a firm within a product line and thereby shortens the duration for incumbents to accumulate markups. Firms further react to the increased risk of replacement by lowering their own-innovation efforts. Both effects compress markup dispersion in the cross section of product lines, thereby improving static efficiency.

The profit tax and research subsidy affect the firm as follows. Following the notation in section 4 production profits in line i are given by

$$\pi_t(\mu_i) = (1-\chi_t)Y_t\left(1-\frac{1}{\mu_{it}}\right),$$

where χ_t denotes the government's tax rate. After-tax profits equal pre-tax profits minus the tax. The firm faces own-innovation and creative destruction research costs of

$$w_t\Gamma\left(x_{ti}, I_{ti}, \mu_{ti}\right) = w_t\left[(1-\upsilon_{It})\mu_{ti}^{-1}\frac{1}{\psi_{Ii}}I_{ti}^{\zeta} + (1-\upsilon_{xt})\frac{1}{\psi_x}x_{ti}^{\zeta}\right],$$

where υ_{It} and υ_{xt} denote the own-innovation and creative destruction specific subsidy rates. I allow for subsidies of both creative destruction and own-innovation costs. This is the same cost formulation as in section 4 with the addition that the firm gets rebated a fraction (υ_{It} or υ_{xt}) of its research costs. This formulation reduces to the model without the government with $\chi_t = \upsilon_{it} = \upsilon_{xt} = 0$. The government satisfies it's budget constraint

$$\chi_t Y_t(1-\Lambda_t) = w_t\int_0^1\left[\upsilon_{It}\mu_i^{-1}\frac{1}{\psi_{Ii}}I_{ti}^{\zeta} + \upsilon_{xt}\frac{1}{\psi_x}x_{ti}^{\zeta}\right]di.$$

The government's budget constraint requires that total tax revenues ($1-\Lambda_t$ is the aggregate profit share) equal total research subsidies paid out (right hand side). For a given tax rate χ_t, I solve for the creative destruction or own-innovation subsidy rate that clears the government's budget constraint jointly with the other equilibrium conditions.

Table 9 shows the effect of a 10% profit tax on the aggregate growth rate and static efficiency under different policies. I consider a subsidy on creative destruction costs, a subsidy on own-innovation costs and a subsidy on both costs (all financed by a 10% profit tax). The first row shows the subsidy on creative destruction. The subsidy that clears the government's budget constraint is a 26% subsidy on the firms' creative destruction costs. The creative destruction subsidy results in a decrease in the aggregate growth rate of 0.2pp. Firm react to the increase in creative destruction by lowering their own-innovation efforts. The increased risk of replacement lowers the incentives for

[18] The government could alternatively subsidize entry costs. Table 15 shows that firm entry contributes very little to aggregate growth. The effects of such policy are therefore small. Instead of taxing profits the government could also tax revenues of the firm. Taxing revenues instead of profits affects firms' incentives for own-innovation and creative destruction differently. The results with a revenue tax are quantitatively different, but qualitatively identical to the results with a profit tax.

firms to accumulate markups. The decrease in own-innovation weighs large on the aggregate growth rate. Own-innovation is the driver of economic growth in the estimated model. In Appendix D, I decompose economic growth into its origins. I show that most economic growth is generated by own-innovation (68%). Creative destruction only accounts for 32%. The fall in own-innovation therefore has a large effect on the aggregate growth rate. The upside of such policy is that it improves static efficiency. The reduction in markup dispersion is equivalent to a permanent 0.41% gain in output (captured by $\Delta\mathcal{M}$). The increase in creative destruction, however, leads to a reallocation of sales shares from the high to the low productivity firms, which decreases aggregate productivity and output permanently by 0.17% (captured by $\Delta\Phi$). The creative destruction policy therefore has contrasting effects on static efficiency. As the effect of reduced markup dispersion outweighs the effect of reduced aggregate productivity, the creative destruction subsidy overall improves static efficiency. In sum, a subsidy on creative destruction costs decreases the aggregate growth rate but improves static efficiency. Trading off economic growth and misallocation, life time utility is 3.7% lower (in permanent consumption equivalents) in the BGP with the creative destruction subsidy than without the subsidy (last column).

Table 9: Effect of innovation subsidies on growth and misallocation

	Subsidy rates		*Growth rate*	*Static efficiency*		*Utility*
	Creative destr.	Own-inno.	Δg	$\Delta\mathcal{M}$	$\Delta\Phi$	$\xi\%$
Profit tax Tax rate = 0.1	0.26		-0.2pp	+0.41%	-0.17%	-3.7%
Profit tax Tax rate = 0.1		0.29	+0.4pp	-0.51%	-0.03%	+5.3%
Profit tax Tax rate = 0.1	0.16	0.16	+0.1pp	+0.09%	-0.14%	+0.3%

Notes: the table shows the effect of different innovation subsidies financed by a 10% profit tax on the agg. growth rate and static efficiency. $\Delta\mathcal{M}$ and $\Delta\Phi$ capture the permanent change in output due to changes in markup dispersion and agg. productivity. ξ denotes the permanent increase in consumption in the BGP without the subsidy that equates lifetime utility of both BGPs.

The second row of Table 9 shows the effects of a subsidy on the own-innovation costs of a firm. A 29% subsidy rate clears the government's budget constraint in equilibrium. The effect on economic growth is strong and positive, raising the aggregate growth rate by 0.4pp. The own-innovation subsidy has a strong effect on the aggregate growth rate, because own-innovation is the main driver of economic growth. The increase in growth does not come without a cost, however. Own-innovation generates economic growth but introduces static efficiency losses through markup dispersion. Relative to the BGP without the own-innovation subsidy, detrended (or permanent) output in the new BGP is 0.51% lower as a result of the increase in markup dispersion (captured by $\Delta\mathcal{M}$). Relative to the creative destruction subsidy, detrended output is almost 1% lower. Overall the positive effect on economic growth outweighs the negative static efficiency effect of the own-innovation subsidy. A permanent increase in consumption of 5.3% in the BGP without the subsidy equates lifetime utility of the BGPs with and without the own-innovation subsidy (last column).

In practice, differentiating between own-innovation and creative destruction costs might be infeasible for a policymaker. As a last counterfactual, I study the effects of a subsidy on the total innovation costs of the firm (own-innovation and creative destruction). Both subsidies in isolation have contrasting effects on economic growth and static efficiency. The effect of a general research subsidy on growth and static efficiency therefore depends on the relative strength of each effect. A general research subsidy that subsidizes creative destruction and own-innovation at an identical rate (last row) has a muted yet overall positive effect on the aggregate growth rate. The positive growth effect of the own-innovation subsidy outweighs the negative effect of the creative destruction subsidy. The subsidy leads to a 0.1pp increase in the aggregate growth rate. The effect on static efficiency is negative but small. A reduction in markup dispersion leads to a permanent 0.09% gain in output, however reallocation of market shares towards less productive firms lowers aggregate productivity by 0.14%. Overall, the difference between lifetime utility of the BGP with and without the general research subsidy is equivalent to a permanent 0.3% increase in consumption (last column).

In the counterfactuals, the own-innovation cost subsidy generates economic growth but lowers static efficiency. The creative destruction cost subsidy improves static efficiency but harms economic growth posing a tradeoff for policymakers to address economic growth and static efficiency.

6 Conclusion

I show in this paper that sales concentration and dispersion of revenue productivity within industries increased in Sweden during 1997–2017. Those trends are in line with recent developments in the US economy. While the trends raise concerns about increasing misallocation and market power at the industry level, I find that firm life cycle growth of sales, revenue productivity, and employment accelerated for firms established after 1997. One exciting avenue for future research is to document whether a similar acceleration of firm life cycle growth can be observed for other countries. Van Vlokhoven (2021) provides suggestive evidence of accelerating sales life cycle growth for US firms in Compustat data and documents a steepening of the profit-age gradient.

To explain the empirical trends, I build a dynamic model of heterogeneous firms with endogenous markups. Firms grow in size by improving their competitors' products through creative destruction, thereby stealing sales shares. On the other hand, firms increase their markups by distancing themselves vertically from their competitors that produce the same product through own-innovation. Sales and markup growth vary systematically across firms. Firms differ with respect to their permanent productivity level, which introduces heterogeneity in their optimal own-innovation and creative destruction rates.

To rationalize the empirical trends, the model asks for a reduction in the own-innovation costs for more productive firms coupled with an increase in the costs of creative destruction (both for incumbent firms and entrants). The changes in innovation costs capture the notion that firms face rising barriers to entering new product markets, whereas improving products in markets where the firm has already established itself as the market leader has become easier. One plausible driver that can simultaneously account for the changes in own-innovation and creative destruction costs is the rising importance of software and networks. Globalization and digitalization gave prominence to such intangible inputs for product planning, development, production, and distribution. The costly acquisition of software and networks deters potential entrants while lowering the cost for product improvements of incumbent firms that have already acquired such infrastructure. Chiavari and Goraya (2020) argue that the rising importance of intangible capital can explain the increase in revenue productivity dispersion and sales concentration in the US. De Ridder (2019) also links the rise in intangible inputs to the recent trends in the US economy. Future work could try to document to what extent the rise in intangible capital facilitates incumbent own-innovation and deters potential entrants.

In response to the changes in the innovation costs, the model predicts a long-run decline in the entry rate, an increase in the aggregate markup, and an increase in the permanent output loss due to markup dispersion from 1.2% to 1.9%. There is a positive side effect to the rise in misallocation. The rise in own-innovation, despite worsening static efficiency, generates economic growth. The acceleration of firm markup growth is associated with an increase in the long-run aggregate growth rate by almost half a percentage point. Aggregate TFP growth and revenue productivity dispersion started to accelerate in the US and the Swedish economy in the mid-1990s, indicating a simultaneous rise in economic growth and misallocation. In fact, I observe that during 1997–2007, a period of high aggregate TFP growth, revenue productivity growth of firms accelerated the fastest.

To address the rise in misallocation, I quantify the effect of different policies on static efficiency. A subsidy on the firms' creative destruction costs that is financed through a profit tax reduces misallocation but decreases the aggregate growth rate as firms lower their own-innovation efforts. The subsidy improves static efficiency at the expense of economic growth. A subsidy on the own-innovation costs of firms instead generates economic growth but increases misallocation. Innovation policies face a tradeoff between economic growth and static efficiency in this model.

I study the effects of permanent changes in innovation costs on the long-run economy. Time will tell if firm growth rates remain permanently high or return to their 1997 levels. In both cases, accounting for transitional dynamics in the model is a possible next step.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N. and Kerr, W. (2018), 'Innovation, reallocation, and growth', *American Economic Review* **108**(11), 3450–91.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J. and Li, H. (2019), A theory of falling growth and rising rents, Technical report, National Bureau of Economic Research.

Aghion, P. and Howitt, P. (1992), 'A Model of Growth through Creative Destruction', *Econometrica* **60**(2), 323–351.

Akcigit, U. and Ates, S. T. (2019), What happened to us business dynamism?, Technical report, National Bureau of Economic Research.

Akcigit, U. and Ates, S. T. (2021), 'Ten facts on declining business dynamism and lessons from endogenous growth theory', *American Economic Journal: Macroeconomics* **13**(1), 257–98.

Akcigit, U. and Kerr, W. R. (2018), 'Growth through heterogeneous innovations', *Journal of Political Economy* **126**(4), 1374–1443.

Andrews, D., Criscuolo, C. and Gal, P. N. (2015), 'Frontier firms, technology diffusion and public policy: Micro evidence from oecd countries'.

Andrews, D., Criscuolo, C. and Gal, P. N. (2016), 'The best versus the rest: the global productivity slowdown, divergence across firms and the role of public policy'.

Atkeson, A. and Burstein, A. (2008), 'Pricing-to-market, trade costs, and international relative prices', *American Economic Review* **98**(5), 1998–2031.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J. (2017), 'Concentrating on the fall of the labor share', *American Economic Review* **107**(5), 180–85.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J. (2020), 'The fall of the labor share and the rise of superstar firms', *The Quarterly Journal of Economics* **135**(2), 645–709.

Bils, M., Klenow, P. J. and Ruane, C. (2021), 'Misallocation or mismeasurement?', *Journal of Monetary Economics* **124**, S39–S56.

Blundell, R., Griffith, R. and Windmeijer, F. (2002), 'Individual effects and dynamics in count data models', *Journal of econometrics* **108**(1), 113–131.

Bornstein, G. (2018), 'Entry and profits in an aging economy: The role of consumer inertia'.

Chatterjee, S. and Eyigungor, B. (2019), 'The firm size and leverage relationship and its implications for entry and concentration in a low interest rate world'.

Chiavari, A. and Goraya, S. (2020), The rise of intangible capital and the macroeconomic implications, Technical report, Working paper.

Crouzet, N. and Eberly, J. C. (2019), Understanding weak capital investment: The role of market concentration and intangibles, Technical report, National Bureau of Economic Research.

De Loecker, J. and Eeckhout, J. (2018), Global market power, Technical report, National Bureau of Economic Research.

De Loecker, J. and Warzynski, F. (2012), 'Markups and firm-level export status', *American economic review* **102**(6), 2437–71.

De Ridder, M. (2019), 'Market power and innovation in the intangible economy'.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2016), 'Declining business dynamism: What we know and the way forward', *American Economic Review* **106**(5), 203–07.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2017), 'Declining dynamism, allocative efficiency, and the productivity slowdown', *American Economic Review* **107**(5), 322–26.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2020), 'Changing business dynamism and productivity: Shocks versus responsiveness', *American Economic Review* **110**(12), 3952–90.

Edmond, C., Midrigan, V. and Xu, D. Y. (2018), How costly are markups?, Technical report, National Bureau of Economic Research.

Elsby, M. W., Hobijn, B. and Şahin, A. (2013), 'The decline of the us labor share', *Brookings Papers on Economic Activity* **2013**(2), 1–63.

Engbom, N. (2020), 'Misallocative growth', *Working Paper* .

Foster, L., Haltiwanger, J. and Syverson, C. (2008), 'Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?', *American Economic Review* **98**(1), 394–425.

Gamber, W. L. (2021), 'Entry, variable markups, and business cycles'.

Garcia-Macia, D., Hsieh, C.-T. and Klenow, P. J. (2019), 'How destructive is innovation?', *Econometrica* **87**(5), 1507–1541.

Gourio, F., Messer, T. and Siemer, M. (2014), A missing generation of firms? aggregate effects of the decline in new business formation, *in* 'CES Ifo Conference'.

Grossman, G. M. and Helpman, E. (1991), 'Quality Ladders in the Theory of Growth', *Review of Economic Studies* **58**(1), 43–61.

Grullon, G., Larkin, Y. and Michaely, R. (2019), 'Are us industries becoming more concentrated?', *Review of Finance* **23**(4), 697–743.

Hall, B. H. and Ziedonis, R. H. (2001), 'The patent paradox revisited: an empirical study of patenting in the us semiconductor industry, 1979-1995', *rand Journal of Economics* pp. 101–128.

Hopenhayn, H., Neira, J. and Singhania, R. (2018), From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share, Technical report, National Bureau of Economic Research.

Hsieh, C.-T. and Klenow, P. J. (2009), 'Misallocation and manufacturing tfp in china and india', *The Quarterly journal of economics* **124**(4), 1403–1448.

Hsieh, C.-T. and Rossi-Hansberg, E. (2019), The industrial revolution in services, Technical report, National Bureau of Economic Research.

Karabarbounis, L. and Neiman, B. (2014), 'The global decline of the labor share', *The Quarterly journal of economics* **129**(1), 61–103.

Karahan, F., Pugsley, B. and Şahin, A. (2019), Demographic origins of the startup deficit, Technical report, National Bureau of Economic Research.

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

Klette, T. J. and Kortum, S. (2004), 'Innovating firms and aggregate innovation', *Journal of political economy* **112**(5), 986–1018.

Lawrence, R. Z. (2015), Recent declines in labor's share in us income: A preliminary neoclassical account, Technical report, National Bureau of Economic Research.

Lentz, R. and Mortensen, D. T. (2008), 'An empirical model of growth through product innovation', *Econometrica* **76**(6), 1317–1373.

Liu, E., Mian, A. and Sufi, A. (2022), 'Low interest rates, market power, and productivity growth', *Econometrica* **90**(1), 193–221.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *econometrica* **71**(6), 1695–1725.

Olmstead-Rumsey, J. (2019), 'Market concentration and the productivity slowdown'.

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

Peters, M. and Walsh, C. (2021), Population growth and firm dynamics, Technical report, National Bureau of Economic Research.

Restuccia, D. and Rogerson, R. (2008), 'Policy distortions and aggregate productivity with heterogeneous establishments', *Review of Economic dynamics* **11**(4), 707–720.

Sandström, M. (2020), *Intangible Capital, Markups and Profits.*

Sterk, V., Sedláček, P. and Pugsley, B. (2021), 'The nature of firm growth', *American Economic Review* **111**(2), 547–79.

Van Reenen, J. (2018), 'Increasing differences between firms: Market power and the macro-economy'.

Van Vlokhoven, H. (2021), 'The rise in profits and fall in firm entry: A tale of the life cycle of profits', *Available at SSRN 3866852* .

Weiss, J. (2019), Intangible investment and market concentration, Technical report, Working paper.

Appendices

A Trends in the Swedish economy

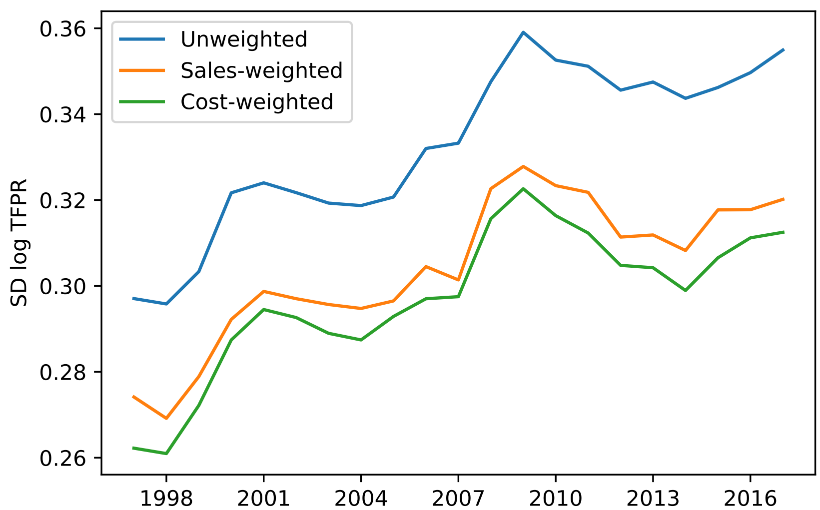
A.1 Rise in $TFPR$ dispersion

In this section I document robustness for the increase in industry $TFPR$ dispersion.

A.1.1 Sales- and cost-weighted standard deviation

Figure 5 shows the sales- and (labor) cost-weighted standard deviation of log $TFPR$ averaged across all 5-digit industries. The unweighted standard deviation is the one shown in the main text.

Figure 5: Weighted standard deviation of log $TFPR$

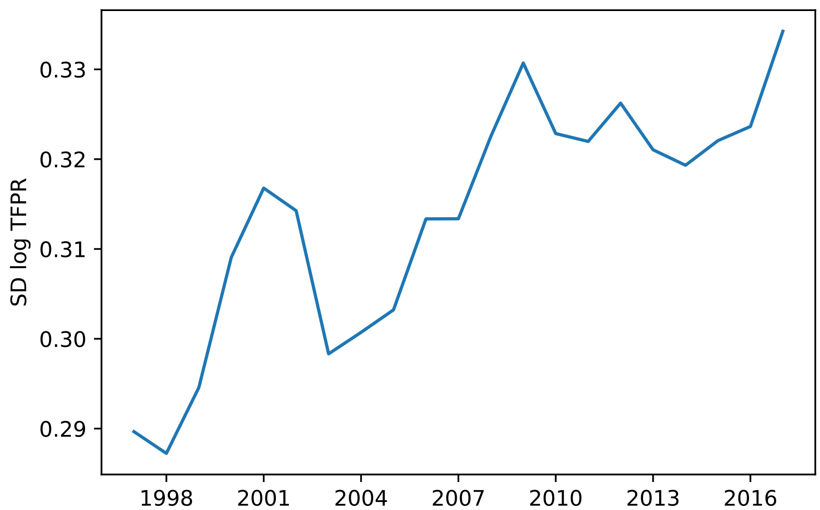


Notes: standard deviation of log $TFPR$ averaged across all (5-digit) industries. Orange and green show the sales- and (labor) cost-weighted standard deviation. $TFPR = VA/(K^{\alpha}(wL)^{1-\alpha})$.

A.1.2 Standard deviation for large industries

Figure 6 shows the standard deviation of log $TFPR$ averaged across industries with at least 50 firms.

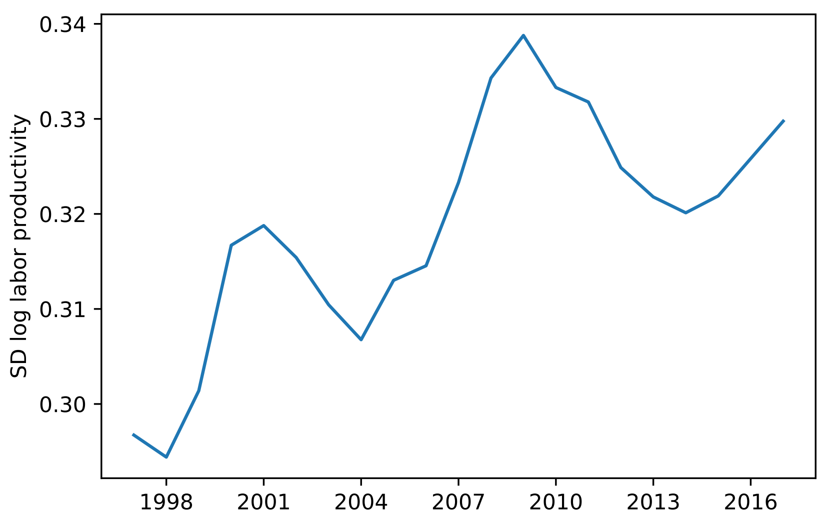
Figure 6: log $TFPR$ standard deviation for large industries



Notes: standard deviation of log $TFPR$ averaged across all (5-digit) industries. Only industries with at least 50 firms included. $TFPR = VA/(K^{\alpha}(wL)^{1-\alpha})$.

A.1.3 Dispersion in labor productivity

Figure 7: Industry labor productivity distribution over time

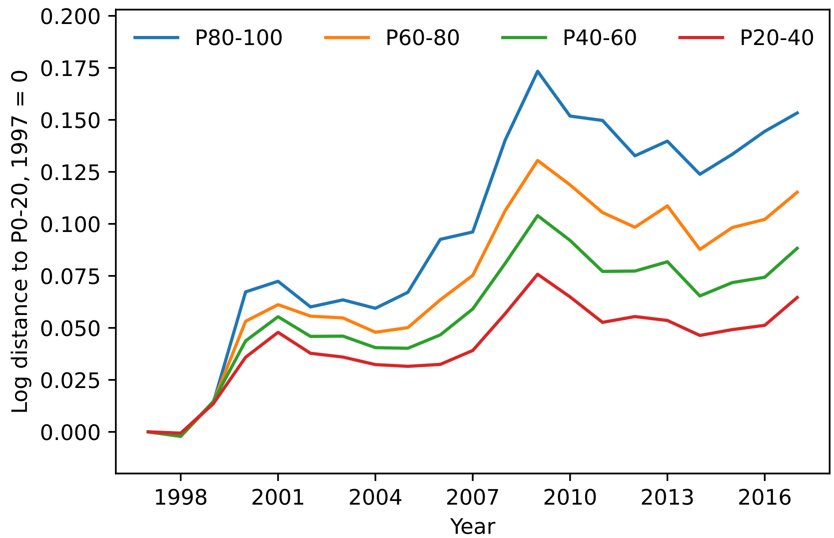


Notes: labor productivity computed as value added over the wage bill of the firm. Standard deviation of log labor productivity averaged across all (5-digit) industries.

A.1.4 Distribution of $TFPR$

Figure 8 shows the average (unweighted) industry log $TFPR$ gap between a given quantile (P80-100, P60-80, P40-60, P20-40) and the bottom quantile (P0-20) of the industry $TFPR$ distribution. The distance to the bottom quantile increased over time for all quantiles. The higher the quantile the further the gap widened.

Figure 8: Industry $TFPR$ distribution over time



Notes: $TFPR = VA/(K^{\alpha}(wL)^{1-\alpha})$. Figure shows the within-industry $\ln TFPR$ gap between a given quantile and the bottom quantile of the $TFPR$ distribution averaged across industries (unweighted). Gap normalized to zero in 1997.

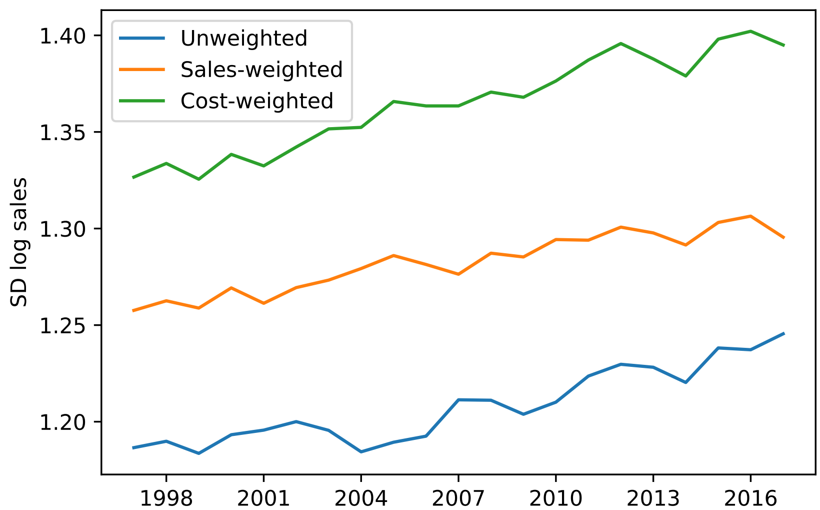
A.2 Rise in sales concentration

This section provides robustness to the reported increase in sales concentration in the main text.

A.2.1 Weighted standard deviations

Figure 9 shows the sales- and (labor) cost-weighted standard deviation of log sales over time.

Figure 9: Weighted standard deviation log sales

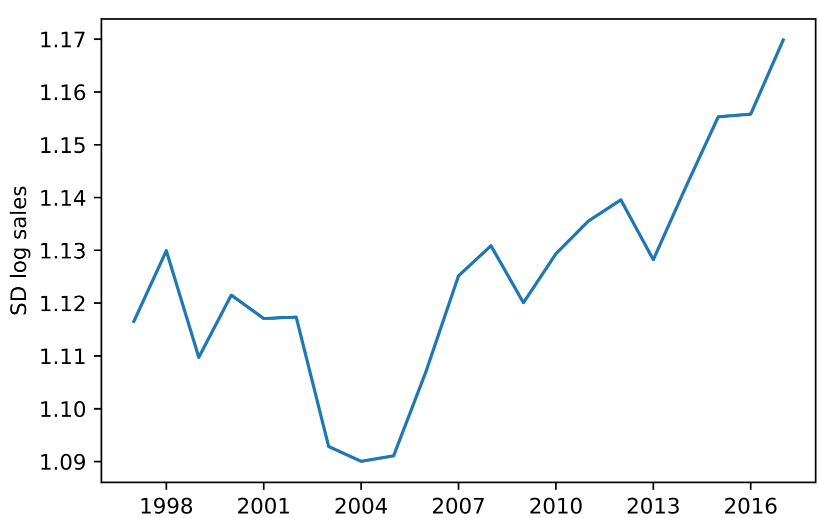


Notes: standard deviation of log sales averaged across all (5-digit) industries. Orange and green show the sales- and (labor) cost-weighted standard deviation.

A.2.2 Standard deviation log sales for large industries

Figure 10 shows the standard deviation of log sales for industries with at least 50 firms over time.

Figure 10: Standard deviation of log sales for large industries

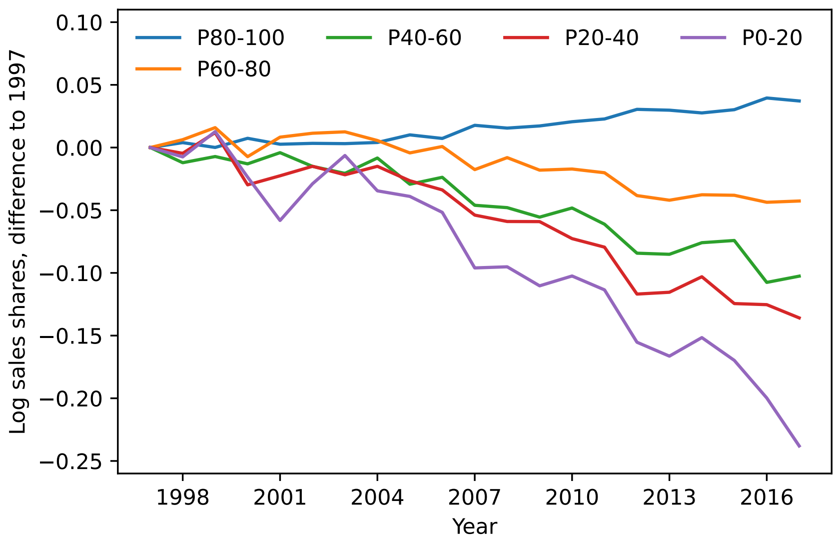


Notes: standard deviation of log sales averaged across all (5-digit) industries. Only industries with at least 50 firms included.

A.2.3 Sales shares by quantile

Figure 11 shows the percentage change in average (unweighted) industry sales shares for different quantiles of the industry sales distribution. Higher quantiles of the sales distribution saw an increase in their average sales share, whereas lower quantiles had their sales shares decline over time, i.e., the industry sales distribution widened.

Figure 11: Changes in sales share by quantile over time

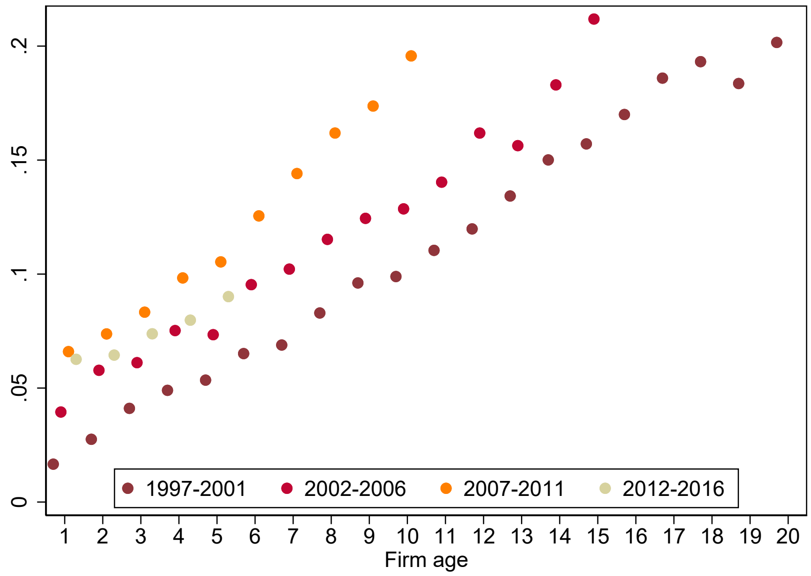


Notes: figure shows the change in a quantile's average (unweighted) industry log sales share. Quantiles refer to quantiles of the within-industry sales distribution.

A.3 Acceleration in firm life cycle growth

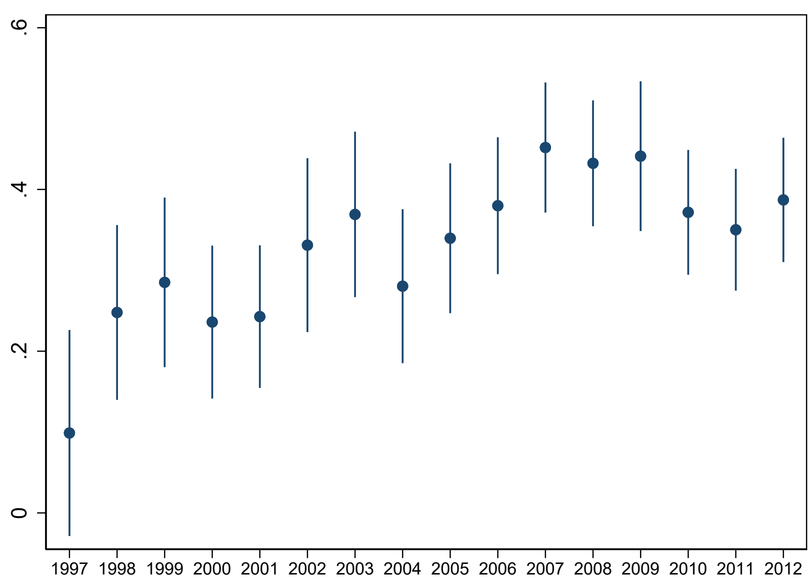
This section provides robustness to the reported increase in firm life cycle growth.

A.3.1 Employment life cycle growth



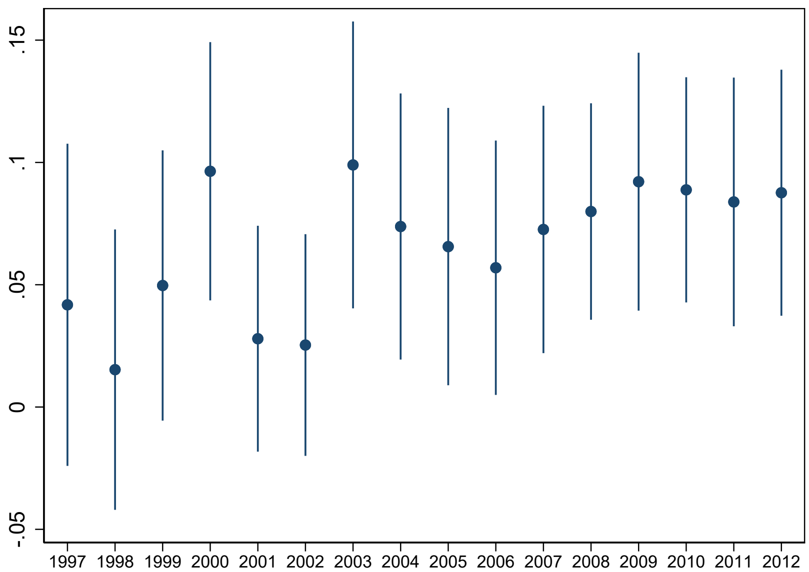
Notes: dots reflect the difference in log employment to age zero (relative to the industry mean). Cohorts pooled as indicated in the legend.

A.3.2 Four year sales growth over time



Notes: coefficients obtained by regressing log sales (net of industry × year FE) on age dummies and a constant (age zero left out) for each cohort separately. Coefficients on the age four dummy shown (with 95% confidence bands), which reflect the change in log sales between age zero and age four relative to the industry mean. The x-axis indicates the cohort.

A.3.3 Four year revenue productivity growth over time



Notes: coefficients obtained by regressing log $TFPR$ (net of industry $\times$ year FE) on age dummies and a constant (age zero left out) for each cohort separately. Coefficients on the age four dummy shown (with 95% confidence bands), which reflect the change in log $TFPR$ between age zero and age four relative to the industry mean. The x-axis indicates the cohort.

B Model

B.1 Solving the dynamic firm problem

The HJB for a high process efficiency firm h reads

$$
\begin{aligned}
& r_t V_t^h(n, [\mu_i]) - \dot{V}_t^h(n, [\mu_i]) = \\
& \qquad \sum_{k=1}^{n} \pi(\mu_k) + \sum_{k=1}^{n} \tau_t \Big[V_t^h\left(n-1, [\mu_i]_{i \neq k}\right) - V_t^h(n, [\mu_i]) \Big] \\
& \qquad + \max_{[x_k, I_k]} \Bigg\{ \sum_{k=1}^{n} I_k \Big[V_t^h\left(n, [[\mu_i]_{i \neq k}, \mu_k \times \lambda]\right) - V_t^h(n, [\mu_i]) \Big] \\
& \qquad + \sum_{k=1}^{n} x_k \Big[S_t V_t^h\left(n+1, [[\mu_i], \lambda]\right) + (1 - S_t) V_t^h\Big(n+1, \Big[[\mu_i], \lambda \times \varphi^h / \varphi^l\Big]\Big) - V_t^h(n, [\mu_i]) \Big] \\
& \qquad - w_t \left[\mu_k^{-1} \frac{1}{\psi_I^h} (I_k)^{\zeta} + \frac{1}{\psi_x} (x_k)^{\zeta} \right] \Bigg\}
\end{aligned}
$$

As in Peters (2020) guess that the value function of the firm consists of a component that is common to all lines and a line specific component[19]

$$V_t^h(n,[\mu_i]) = V_{t,P}^h(n) + \sum_{k=1}^n V_{t,M}^h(\mu_k)$$

so that

$$\begin{aligned}
\dot{V}_t^h(n,[\mu_i]) &= \dot{V}_{t,P}^h(n) + \sum_{k=1}^n \dot{V}_{t,M}^h(\mu_k)\\
V_t^h\left(n-1,[\mu_i]_{i\neq k}\right) - V_t^h(n,[\mu_i]) &= V_{t,P}^h(n-1) - V_{t,P}^h(n) - V_{t,M}^h(\mu_k)\\
V_t^h\left(n,[[\mu_i]_{i\neq k},\mu_k\times\lambda]\right) - V_t^h(n,[\mu_i]) &= V_{t,M}^h(\mu_k\times\lambda) - V_{t,M}^h(\mu_k)
\end{aligned}$$

and

$$\begin{aligned}
&S_t V_t^h\left(n+1,[[\mu_i],\lambda]\right) + (1-S_t)V_t^h\left(n+1,\left[[\mu_i],\lambda\times\varphi^h/\varphi^l\right]\right) - V_t^h(n,[\mu_i] = \\
&S_t\left(V_{t,P}^h(n+1) + \sum_{k=1}^n V_{t,M}^h(\mu_k) + V_{t,M}^h(\lambda)\right) + (1-S_t)\left(V_{t,P}^h(n+1) + \sum_{k=1}^n V_{t,M}^h(\mu_k) + V_{t,M}^h(\lambda\times\varphi^h/\varphi^l)\right)\\
&-V_{t,P}^h(n) - \sum_{k=1}^n V_{t,M}^h(\mu_k) = \\
&V_{t,P}^h(n+1) - V_{t,P}^h(n) + S_t V_{t,M}^h(\lambda) + (1-S_t)V_{t,M}^h(\lambda\times\varphi^h/\varphi^l)
\end{aligned}$$

Substituting the guess into the HJB

$$\begin{aligned}
r_t\Big[V_{t,P}^h(n) + \sum_{k=1}^n V_{t,M}^h(\mu_k)\Big] - \dot{V}_{t,P}^h(n) - \sum_{k=1}^n \dot{V}_{t,M}^h(\mu_k) = &\\
&\sum_{k=1}^n \pi(\mu_k) + \sum_{k=1}^n \tau_t\Big[V_{t,P}^h(n-1) - V_{t,P}^h(n) - V_{t,M}^h(\mu_k)\Big]\\
&+\max_{[x_k,I_k]}\Bigg\{\sum_{k=1}^n I_k\Big[V_{t,M}^h(\mu_k\times\lambda) - V_{t,M}^h(\mu_k)\Big]\\
&+\sum_{k=1}^n x_k\Big[V_{t,P}^h(n+1) - V_{t,P}^h(n) + S_t V_{t,M}^h(\lambda) + (1-S_t)V_{t,M}^h(\lambda\times\varphi^h/\varphi^l)\Big]\\
&-w_t\left[\sum_{k=1}^n \mu_k^{-1}\frac{1}{\psi_I^h}(I_k)^\zeta + \frac{1}{\psi_x}(x_k)^\zeta\right]\Bigg\}
\end{aligned}$$

and rearranging terms

$$\begin{aligned}
&r_t V_{t,P}^h(n) - \dot{V}_{t,P}^h(n) + \sum_{k=1}^n\Big[r_t V_{t,M}^h(\mu_k) - \dot{V}_{t,M}^h(\mu_k)\Big] = \\
&\quad\sum_{k=1}^n\left\{\pi(\mu_k) - \tau_t V_{t,M}^h(\mu_k) + \max_{I_k}\left\{I_k\Big[V_{t,M}^h(\mu_k\times\lambda) - V_{t,M}^h(\mu_k)\Big] - w_t\mu_k^{-1}\frac{1}{\psi_I^h}(I_k)^\zeta\right\}\right\}\\
&\quad+\sum_{k=1}^n \tau_t\Big[V_{t,P}^h(n-1) - V_{t,P}^h(n)\Big]\\
&\quad+\max_{[x_k]}\left\{\sum_{k=1}^n x_k\Big[V_{t,P}^h(n+1) - V_{t,P}^h(n) + S_t V_{t,M}^h(\lambda) + (1-S_t)V_{t,M}^h(\lambda\times\varphi^h/\varphi^l)\Big] - w_t\frac{1}{\psi_x}(x_k)^\zeta\right\}
\end{aligned}$$

[19]The notation follows Peters (2020) where possible.

First solve

$$r_t V^h_{t,M}(\mu_i) - \dot{V}^h_{t,M}(\mu_i) = \pi(\mu_i) - \tau_t V^h_{t,M}(\mu_i) \\ + \max_{I_i} \Big\{ I_i \Big[V^h_{t,M}(\mu_i \times \lambda) - V^h_{t,M}(\mu_i) \Big] - w_t \mu_i^{-1} \frac{1}{\psi^h_I} (I_i)^\zeta \Big\}. \tag{28}$$

Once we know $V^h_{t,M}$, we can solve for $V^h_{t,P}$ in

$$r_t V^h_{t,P}(n) - \dot{V}^h_{t,P}(n) = \sum_{k=1}^{n} \tau_t \Big[V^h_{t,P}(n-1) - V^h_{t,P}(n) \Big] \\ + \max_{[x_k]} \Big\{ \sum_{k=1}^{n} x_k \Big[V^h_{t,P}(n+1) - V^h_{t,P}(n) + S_t V^h_{t,M}(\lambda) + (1 - S_t) V^h_{t,M}(\lambda \times \varphi^h / \varphi^l) \Big] - w_t \frac{1}{\psi_x} (x_k)^\zeta \Big\} \tag{29}$$

Assume (and verified below) that in steady-state $V^h_{t,P}$ and $V^h_{t,M}$ grow at the constant rate g such that

$$\dot{V}^h_{t,P}(n) = g V^h_{t,P}(n)$$
$$\dot{V}^h_{t,M}(\mu_i) = g V^h_{t,M}(\mu_i)$$

In steady-state we can then write eq. (28) as

$$(r - g + \tau) V^h_{t,M}(\mu_i) = \pi(\mu_i) + \max_{I_i} \Big\{ I_i \Big[V^h_{t,M}(\mu_i \times \lambda) - V^h_{t,M}(\mu_i) \Big] - w_t \mu_i^{-1} \frac{1}{\psi^h_I} (I_i)^\zeta \Big\}. \tag{30}$$

Guess that[20]

$$V^h_{t,M}(\mu_i) = \kappa_t - \alpha_t \mu_i^{-1}$$

so that

$$V^h_{t,M}(\mu_i \times \lambda) - V^h_{t,M}(\mu_i) = \alpha_t \mu_i^{-1} \left(1 - \frac{1}{\lambda} \right).$$

The FOC for I_i then reads

$$\alpha_t \mu_i^{-1} \left(1 - \frac{1}{\lambda} \right) = w_t \mu^{-1} \frac{1}{\psi^h_I} \zeta \left(I_i \right)^{\zeta - 1}.$$

Rearranging yields

$$\left(\frac{\alpha_t}{w_t} \left(1 - \frac{1}{\lambda} \right) \frac{\psi^h_I}{\zeta} \right)^{\frac{1}{\zeta - 1}} = I_i. \tag{31}$$

It will turn out that α_t / w_t is constant such that I_i is time independent. Using the guess for the value function, the FOC, and the Euler equation $\rho = r - g$ in eq. (30) we get

$$(\rho + \tau) \left[\kappa_t - \alpha_t \mu_i^{-1} \right] = Y_t \left(1 - \frac{1}{\mu_i} \right) + w_t \mu^{-1} \frac{1}{\psi^h_I} \zeta \left(I_i \right)^\zeta - w_t \mu_i^{-1} \frac{1}{\psi^h_I} (I_i)^\zeta$$
$$= Y_t \left(1 - \frac{1}{\mu_i} \right) + \frac{\zeta - 1}{\psi^h_I} w_t \mu_i^{-1} (I_i)^\zeta.$$

[20] It will turn out from eq. (31) that $\alpha_t > 0$, otherwise I_i would not be positive, such that $V^h_{t,M}(\mu_i)$ is increasing in μ_i.

Matching coefficients we obtain

$$\begin{aligned}(\rho+\tau)\kappa_t =& Y_t \\ \Leftrightarrow \kappa_t =& \frac{Y_t}{\rho+\tau}\end{aligned}$$

and

$$\begin{aligned}-\alpha_t\mu_i^{-1} =& \frac{-Y_t\mu_i^{-1} + \frac{\zeta-1}{\psi_I^h} w_t\mu_i^{-1}(I_i)^\zeta}{\rho+\tau} \\ \Leftrightarrow \alpha_t =& \frac{Y_t - \frac{\zeta-1}{\psi_I^h}(I_i)^\zeta w_t}{\rho+\tau}.\end{aligned}$$

This confirms that α_t/w_t is indeed time independent. The value function is

$$\begin{aligned}V_{t,M}^h(\mu_i) =& \kappa_t - \alpha_t\mu_i^{-1} \\ =& \frac{Y_t}{\rho+\tau} - \frac{Y_t\mu_i^{-1} - \frac{\zeta-1}{\psi_I^h}(I_i)^\zeta w_t\mu_i^{-1}}{\rho+\tau} \\ =& \frac{\pi(\mu_i) + \frac{\zeta-1}{\psi_I^h}(I_i)^\zeta w_t\mu_i^{-1}}{\rho+\tau}\end{aligned}$$

Inserting α into the optimality condition, I_i solves

$$I_i = \left(\left(\frac{Y_t}{w_t} - \frac{\zeta-1}{\psi_I^h}(I_i)^\zeta\right)\left(1-\frac{1}{\lambda}\right)\frac{\psi_I^h}{\zeta(\rho+\tau)}\right)^{\frac{1}{\zeta-1}}.$$

Own-innovation rates are not product line specific, but just specific to the productivity type of the producer: $I_i = I^h$. They are further time invariant. Optimality condition for high productivity own-innovation rate:

$$I^h = \left(\left(\frac{Y_t}{w_t} - \frac{\zeta-1}{\psi_I^h}(I^h)^\zeta\right)\left(1-\frac{1}{\lambda}\right)\frac{\psi_I^h}{\zeta(\rho+\tau)}\right)^{\frac{1}{\zeta-1}}$$

Equivalently the optimality condition for the own-innovation rate of low productivity firms:

$$I^l = \left(\left(\frac{Y_t}{w_t} - \frac{\zeta-1}{\psi_I^l}(I^l)^\zeta\right)\left(1-\frac{1}{\lambda}\right)\frac{\psi_I^l}{\zeta(\rho+\tau)}\right)^{\frac{1}{\zeta-1}}.$$

With this at hand we can turn back to the differential equation for $V_{t,P}^h(n)$ in eq. (29).

$$\begin{aligned}r_tV_{t,P}^h(n) - \dot{V}_{t,P}^h(n) = \sum_{k=1}^n \tau_t\Big[V_{t,P}^h(n-1) - V_{t,P}^h(n)\Big] \\ + \max_{[x_k]}\Big\{\sum_{k=1}^n x_k\Big[V_{t,P}^h(n+1) - V_{t,P}^h(n) + S_tV_{t,M}^h(\lambda) + (1-S_t)V_{t,M}^h(\lambda\times\varphi^h/\varphi^l)\Big] - w_t\frac{1}{\psi_x}(x_k)^\zeta\Big\}\end{aligned}$$

In addition to the guess that $V^h_{t,P}(n)$ grows at rate g, conjecture that $V^h_{t,P}(n) = n \times v^h_t$. Combined with the Euler we get

$$(\rho+\tau)nv^h_t = \max_{[x_k]} \Big\{ \sum_{k=1}^{n} x_k \Big[v^h_t + S_t V^h_{t,M}(\lambda) + (1-S_t)V^h_{t,M}(\lambda \times \varphi^h/\varphi^l) \Big] - w_t \frac{1}{\psi_x}(x_k)^{\zeta} \Big\}. \tag{32}$$

The optimality condition for x_k is given by

$$v^h_t + S_t V^h_{t,M}(\lambda) + (1-S_t)V^h_{t,M}(\lambda \times \varphi^h/\varphi^l) = w_t \frac{\zeta}{\psi_x}(x_k)^{\zeta-1}. \tag{33}$$

Several observations are noteworthy. First, the FOC shows that optimal expansion rates are independent of quality and productivity gaps in line k. Second, the optimal expansion rate x_k depends on the own productivity type through the value function and the expansion cost. We can hence drop the item indexation: $x_k = x^h$. This is intuitive since for expansion a random new line will be drawn. The expected value of acquiring a new line depends on the own productivity. Since productivity is firm specific this implies that expansion rates across lines within the same firm are identical. Third, $v_t, V^h_{t,M}, w_t$ all grow at the same rate g, which implies that expansion rates are constant over time. We can hence write eq. (32) as

$$(\rho+\tau)nv^h_t = nw_t \frac{\zeta}{\psi_x}(x^h)^{\zeta} - nw_t \frac{1}{\psi_x}(x^h)^{\zeta}.$$

or

$$v^h_t = \frac{1}{(\rho+\tau)} \frac{\zeta-1}{\psi_x}(x^h)^{\zeta} w_t.$$

Gathering all terms the value function is given by

$$\begin{aligned}
V^h_t(n,[\mu_i]) =& V^h_{t,P}(n) + \sum_{k=1}^{n} V^h_{t,M}(\mu_k) \\
=& nv^h_t + \sum_{k=1}^{n} V^h_{t,M}(\mu_k) \\
=& n\frac{1}{(\rho+\tau)} \frac{\zeta-1}{\psi_x}(x^h)^{\zeta} w_t + \sum_{k=1}^{n} \frac{\pi(\mu_k) + \frac{\zeta-1}{\psi^h_I}(I^h)^{\zeta} w_t \mu_k^{-1}}{\rho+\tau}
\end{aligned}$$

Using the expression for v^h_t in the optimality condition for the expansion rate write eq. (33) as

$$\begin{aligned}
\frac{1}{(\rho+\tau)} \frac{\zeta-1}{\psi_x}(x^h)^{\zeta} w_t + S_t \frac{\pi(\lambda) + \frac{\zeta-1}{\psi^h_I}(I^h)^{\zeta} w_t \lambda^{-1}}{\rho+\tau} + (1-S_t) \frac{\pi(\lambda \times \varphi^h/\varphi^l) + \frac{\zeta-1}{\psi^h_I}(I^h)^{\zeta} w_t \lambda^{-1} \frac{\varphi^l}{\varphi^h}}{\rho+\tau} \\
= w_t \frac{\zeta}{\psi_x}(x^h)^{\zeta-1}.
\end{aligned}$$

Simplifying gives

$$\begin{aligned}
\frac{\zeta-1}{\psi_x}(x^h)^{\zeta} + S_t \left(\pi(\lambda)/w_t + \frac{\zeta-1}{\psi^h_I}(I^h)^{\zeta}\lambda^{-1} \right) + (1-S_t) \left(\pi(\lambda \times \varphi^h/\varphi^l)/w_t + \frac{\zeta-1}{\psi^h_I}(I^h)^{\zeta}\lambda^{-1}\frac{\varphi^l}{\varphi^h} \right) \\
= (\rho+\tau)\frac{\zeta}{\psi_x}(x^h)^{\zeta-1}.
\end{aligned}$$

Inserting the profit function

$$\frac{\zeta-1}{\psi_x}(x^h)^\zeta + S_t\left(\frac{Y_t}{w_t}\left(1-\frac{1}{\lambda}\right)+\frac{\zeta-1}{\psi_I^h}(I^h)^\zeta\lambda^{-1}\right)+(1-S_t)\left(\frac{Y_t}{w_t}\left(1-\frac{\varphi^l}{\varphi^h}\frac{1}{\lambda}\right)+\frac{\zeta-1}{\psi_I^h}(I^h)^\zeta\lambda^{-1}\frac{\varphi^l}{\varphi^h}\right)$$
$$=(\rho+\tau)\frac{\zeta}{\psi_x}(x^h)^{\zeta-1}.$$

The optimality condition for the expansion rate of the low productivity type reads

$$\frac{\zeta-1}{\psi_x}(x^l)^\zeta + S_t\left(\frac{Y_t}{w_t}\left(1-\frac{1}{\lambda}\frac{\varphi^h}{\varphi^l}\right)+\frac{\zeta-1}{\psi_I^l}(I^l)^\zeta\lambda^{-1}\frac{\varphi^h}{\varphi^l}\right)+(1-S_t)\left(\frac{Y_t}{w_t}\left(1-\frac{1}{\lambda}\right)+\frac{\zeta-1}{\psi_I^l}(I^l)^\zeta\lambda^{-1}\right)$$
$$=(\rho+\tau)\frac{\zeta}{\psi_x}(x^l)^{\zeta-1}.$$

B.2 Solving for the steady state equilibrium

In the model there are the eight unknown variables $x^h, x^l, I^h, I^l, z, \tau, \frac{Y_t}{w_t}, S$ and the markup distribution $\nu()$ in eight equations plus the system of differential equations characterizing $\nu()$.

Optimality condition for high productivity own-innovation rate

$$I^h=\left(\left(\frac{Y_t}{w_t}-\frac{\zeta-1}{\psi_I^h}(I^h)^\zeta\right)\left(1-\frac{1}{\lambda}\right)\frac{\psi_I^h}{\zeta(\rho+\tau)}\right)^{\frac{1}{\zeta-1}}$$

Optimality condition for low productivity own-innovation rate

$$I^l=\left(\left(\frac{Y_t}{w_t}-\frac{\zeta-1}{\psi_I^l}(I^l)^\zeta\right)\left(1-\frac{1}{\lambda}\right)\frac{\psi_I^l}{\zeta(\rho+\tau)}\right)^{\frac{1}{\zeta-1}}$$

Optimality condition for high productivity expansion rate

$$\frac{\zeta-1}{\psi_x}(x^h)^\zeta + S\left(\frac{Y_t}{w_t}\left(1-\frac{1}{\lambda}\right)+\frac{\zeta-1}{\psi_I^h}(I^h)^\zeta\lambda^{-1}\right)+(1-S)\left(\frac{Y_t}{w_t}\left(1-\frac{\varphi^l}{\varphi^h}\frac{1}{\lambda}\right)+\frac{\zeta-1}{\psi_I^h}(I^h)^\zeta\lambda^{-1}\frac{\varphi^l}{\varphi^h}\right)$$
$$=(\rho+\tau)\frac{\zeta}{\psi_x}(x^h)^{\zeta-1}$$

Optimality condition for low productivity expansion rate

$$\frac{\zeta-1}{\psi_x}(x^l)^\zeta + S\left(\frac{Y_t}{w_t}\left(1-\frac{1}{\lambda}\frac{\varphi^h}{\varphi^l}\right)+\frac{\zeta-1}{\psi_I^l}(I^l)^\zeta\lambda^{-1}\frac{\varphi^h}{\varphi^l}\right)+(1-S)\left(\frac{Y_t}{w_t}\left(1-\frac{1}{\lambda}\right)+\frac{\zeta-1}{\psi_I^l}(I^l)^\zeta\lambda^{-1}\right)$$
$$=(\rho+\tau)\frac{\zeta}{\psi_x}(x^l)^{\zeta-1}$$

Free entry condition

$$p^h\Big(SV_t^h(1,\lambda)+(1-S)V_t^h(1,\lambda\times\varphi^h/\varphi^l)\Big)+(1-p^h)\Big(SV_t^l(1,\lambda\times\varphi^l/\varphi^h)+(1-S)V_t^l(1,\lambda)\Big)=\frac{1}{\psi_z}w_t,$$

where

$$V_t^d(1,\mu)=\frac{1}{(\rho+\tau)}\frac{\zeta-1}{\psi_x}(x^d)^\zeta w_t+\frac{Y_t\left(1-\frac{1}{\mu}\right)+\frac{\zeta-1}{\psi_I^d}(I^d)^\zeta w_t\mu^{-1}}{\rho+\tau}$$

Labor market clearing condition

$$1=\frac{Y_t}{w_t}\sum_{\frac{\varphi_j}{\varphi_{j'}}}\sum_i \frac{1}{\lambda^{i\frac{\varphi_j}{\varphi_{j'}}}}\nu\left(i,\frac{\varphi_j}{\varphi_{j'}}\right)+\sum_{\frac{\varphi_j}{\varphi_{j'}}}\frac{1}{\psi_I^j}(I^j)^\zeta\sum_i \frac{1}{\lambda^{i\frac{\varphi_j}{\varphi_{j'}}}}\nu\left(i,\frac{\varphi_j}{\varphi_{j'}}\right)+S\frac{1}{\psi_x}(x^h)^\zeta+(1-S)\frac{1}{\psi_x}(x^l)^\zeta+\frac{z}{\psi_z}$$

Creative destruction

$$\tau=z+Sx^h+(1-S)x^l$$

Share of high productivity type

$$S=\sum_{i=1}^{\infty}\left[\nu\left(i,\frac{\varphi^h}{\varphi^h}\right)+\nu\left(i,\frac{\varphi^h}{\varphi^l}\right)\right],$$

where ν, the stationary distribution of quality and productivity gaps, is characterized by

$$0=\dot{\nu}\left(\Delta,\frac{\varphi_j}{\varphi_{j'}}\right)=I^l\nu\left(\Delta-1,\frac{\varphi_j}{\varphi_{j'}}\right)-\nu\left(\Delta,\frac{\varphi_j}{\varphi_{j'}}\right)(I^l+\tau)\quad\text{for}\quad\Delta\geq 2,\frac{\varphi_j}{\varphi_{j'}}\in\left\{\frac{\varphi^l}{\varphi^h},\frac{\varphi^l}{\varphi^l}\right\}$$
$$0=\dot{\nu}\left(\Delta,\frac{\varphi_j}{\varphi_{j'}}\right)=I^h\nu\left(\Delta-1,\frac{\varphi_j}{\varphi_{j'}}\right)-\nu\left(\Delta,\frac{\varphi_j}{\varphi_{j'}}\right)(I^h+\tau)\quad\text{for}\quad\Delta\geq 2,\frac{\varphi_j}{\varphi_{j'}}\in\left\{\frac{\varphi^h}{\varphi^h},\frac{\varphi^h}{\varphi^l}\right\}$$

and for the case of a unitary quality gap

$$0=\dot{\nu}\left(1,\frac{\varphi^l}{\varphi^h}\right)=(1-S)x^lS+z_t(1-p^h)S-\nu\left(1,\frac{\varphi^l}{\varphi^h}\right)(I^l+\tau)$$
$$0=\dot{\nu}\left(1,\frac{\varphi^l}{\varphi^l}\right)=(1-S)x^l(1-S)+z_t(1-p^h)(1-S)-\nu\left(1,\frac{\varphi^l}{\varphi^l}\right)(I^l+\tau)$$
$$0=\dot{\nu}\left(1,\frac{\varphi^h}{\varphi^h}\right)=Sx^hS+z_tp^hS-\nu\left(1,\frac{\varphi^h}{\varphi^h}\right)(I^h+\tau)$$
$$0=\dot{\nu}\left(1,\frac{\varphi^h}{\varphi^l}\right)=Sx^h(1-S)+z_tp^h(1-S)-\nu\left(1,\frac{\varphi^h}{\varphi^l}\right)(I^h+\tau)$$

To simplify the system of equations, first rewrite the creative destruction equation

$$z=(\tau-Sx^h-(1-S)x^l)$$

such that z can be substituted out from the remaining equations. Second, as derived in the main text, from the differential equations characterizing the distribution of quality and productivity gaps in steady-state we obtain for the share of high productivity incumbents in the economy

$$\begin{aligned}S=&S_{\varphi^h,\varphi^h}+S_{\varphi^h,\varphi^l}\\=&\frac{Sx^h+zp^h}{\tau},\end{aligned}$$

Third, the optimality conditions for expansion rates (multiplied by p^h and $(1-p^h)$) and the free entry condition together imply

$$\frac{1}{\psi_x}p^h(x^h)^{\zeta-1}+\frac{1}{\psi_x}(1-p^h)(x^l)^{\zeta-1}=\frac{1}{\psi_z\zeta}.$$

The system of equilibrium conditions can hence be reduced to:

Optimality condition for high productivity own-innovation rate

$$I^h = \left(\left(\frac{Y_t}{w_t} \psi_I^h - (\zeta - 1)(I^h)^\zeta \right) \frac{\left(1 - \frac{1}{\lambda}\right)}{\zeta(\rho + \tau)} \right)^{\frac{1}{\zeta - 1}}$$

Optimality condition for low productivity own-innovation rate

$$I^l = \left(\left(\frac{Y_t}{w_t} \psi_I^l - (\zeta - 1)(I^l)^\zeta \right) \frac{\left(1 - \frac{1}{\lambda}\right)}{\zeta(\rho + \tau)} \right)^{\frac{1}{\zeta - 1}}$$

Optimality condition for high productivity expansion rate

$$\frac{\zeta - 1}{\psi_x}(x^h)^\zeta + S\left(\frac{Y_t}{w_t}\left(1 - \frac{1}{\lambda}\right) + (\zeta - 1)(I^h)^\zeta \lambda^{-1} \frac{1}{\psi_I^h}\right) + (1 - S)\left(\frac{Y_t}{w_t}\left(1 - \frac{\varphi^l}{\varphi^h}\frac{1}{\lambda}\right) + (\zeta - 1)(I^h)^\zeta \lambda^{-1} \frac{\varphi^l}{\psi_I^h \varphi^h}\right)$$
$$= (\rho + \tau)\frac{\zeta}{\psi_x}(x^h)^{\zeta - 1}$$

Optimality condition for low productivity expansion rate

$$\frac{\zeta - 1}{\psi_x}(x^l)^\zeta + S\left(\frac{Y_t}{w_t}\left(1 - \frac{1}{\lambda}\frac{\varphi^h}{\varphi^l}\right) + (\zeta - 1)(I^l)^\zeta \lambda^{-1} \frac{\varphi^h}{\psi_I^l \varphi^l}\right) + (1 - S)\left(\frac{Y_t}{w_t}\left(1 - \frac{1}{\lambda}\right) + (\zeta - 1)(I^l)^\zeta \lambda^{-1} \frac{1}{\psi_I^l}\right)$$
$$= (\rho + \tau)\frac{\zeta}{\psi_x}(x^l)^{\zeta - 1}$$

Free entry

$$p^h \frac{(x^h)^{\zeta - 1}}{\psi_x} + (1 - p^h)\frac{(x^l)^{\zeta - 1}}{\psi_x} = \frac{1}{\psi_z \zeta}$$

Labor market clearing condition

$$1 = \frac{Y_t}{w_t}\Lambda + \Lambda_I + S\frac{1}{\psi_x}(x^h)^\zeta + (1 - S)\frac{1}{\psi_x}(x^l)^\zeta + \frac{\tau - Sx^h - (1 - S)x^l}{\psi_z},$$

where[21]

$$\Lambda = \sum_{k \in \{h,l\}} \frac{\theta_k}{\theta_k + 1} \sum_{n \in \{h,l\}} \frac{1}{\varphi_k / \varphi_n} S_{\varphi_k, \varphi_n}$$
$$\Lambda_I = \sum_{k \in \{h,l\}} \frac{1}{\psi_I^k}(I^k)^\zeta \frac{\theta_k}{\theta_k + 1} \sum_{n \in \{h,l\}} \frac{1}{\varphi_k / \varphi_n} S_{\varphi_k, \varphi_n}$$
$$\theta_k = \frac{\ln(I^k + \tau) - \ln(I^k)}{\ln(\lambda)}$$

Share of high productivity type

$$S = \frac{Sx^h + (\tau - Sx^h - (1 - S)x^l)p^h}{\tau}$$

This constitutes a system of seven equations in seven unknowns $(x^h, x^l, I^h, I^l, \tau, \frac{Y_t}{w_t}, S)$, which I solve using a root finder.

[21] For the derivation of Λ I assume a continuous distribution of quality gaps.

B.3 Deriving Λ and $\mathcal{M}$

From the two-dimensional distribution of quality and productivity gaps in the balanced growth path it follows that

$$\Pr\left(\Delta \leq d, \frac{\varphi^l}{\varphi^h}\right) = \sum_{i=1}^{d} \nu\left(i, \frac{\varphi^l}{\varphi^h}\right) = S_{\varphi^l,\varphi^h}\left(1 - \left(\frac{I^l}{I^l+\tau}\right)^d\right)$$

$$\Pr\left(\Delta \leq d, \frac{\varphi^l}{\varphi^l}\right) = \sum_{i=1}^{d} \nu\left(i, \frac{\varphi^l}{\varphi^l}\right) = S_{\varphi^l,\varphi^l}\left(1 - \left(\frac{I^l}{I^l+\tau}\right)^d\right)$$

$$\Pr\left(\Delta \leq d, \frac{\varphi^h}{\varphi^h}\right) = \sum_{i=1}^{d} \nu\left(i, \frac{\varphi^h}{\varphi^h}\right) = S_{\varphi^h,\varphi^h}\left(1 - \left(\frac{I^h}{I^h+\tau}\right)^d\right)$$

$$\Pr\left(\Delta \leq d, \frac{\varphi^h}{\varphi^l}\right) = \sum_{i=1}^{d} \nu\left(i, \frac{\varphi^h}{\varphi^l}\right) = S_{\varphi^h,\varphi^l}\left(1 - \left(\frac{I^h}{I^h+\tau}\right)^d\right).$$

Focusing on product lines where a low productivity incumbent faces a high productivity second-best firm:

$$\begin{aligned}
P\left(\frac{\varphi^l}{\varphi^h}, \Delta \leq d\right) =& S_{\varphi^l,\varphi^h}\left(1 - \left(\frac{I^l}{I^l+\tau}\right)^d\right) \\
=& S_{\varphi^l,\varphi^h}\left(1 - e^{\ln\left(\left(\frac{I^l}{I^l+\tau}\right)^d\right)}\right) \\
=& S_{\varphi^l,\varphi^h}\left(1 - e^{-d[\ln(I^l+\tau)-\ln(I^l)]}\right)
\end{aligned}$$

and

$$\begin{aligned}
P\left(\frac{\varphi^l}{\varphi^h}, \ln\left(\lambda^\Delta\right) \leq d\right) =& P\left(\frac{\varphi^l}{\varphi^h}, \Delta \ln(\lambda) \leq d\right) \\
=& P\left(\frac{\varphi^l}{\varphi^h}, \Delta \leq \frac{d}{\ln(\lambda)}\right) \\
=& S_{\varphi^l,\varphi^h}\left(1 - e^{-\frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}d}\right)
\end{aligned}$$

Conditional on the productivity gap, $\ln\left(\lambda^\Delta\right)$ is exponentially distributed with parameter $\frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}$. Further

$$\begin{aligned}
P\left(\frac{\varphi^l}{\varphi^h}, \lambda^\Delta \leq d\right) =& P\left(\frac{\varphi^l}{\varphi^h}, \Delta \leq \frac{\ln(d)}{\ln(\lambda)}\right) \\
=& S_{\varphi^l,\varphi^h}\left(1 - e^{-\frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}\ln(d)}\right) \\
=& S_{\varphi^l,\varphi^h}\left(1 - d^{-\frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}}\right)
\end{aligned}$$

Conditional on the productivity gap, quality gaps follow a Pareto distribution with parameter $\frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}$. Denote $\theta_l = \frac{\ln(I^l+\tau)-\ln(I^l)}{\ln(\lambda)}$. We then have

$$P\left(\frac{\varphi^l}{\varphi^h}, \lambda^\Delta \leq m\right) = S_{\varphi^l,\varphi^h}\left(1 - m^{-\theta_l}\right).$$

Doing the same steps for lines with different productivity gaps, the aggregate labor share can be computed as

$$\begin{aligned}
\Lambda &= \sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\int_1^\infty \frac{1}{\varphi_k/\varphi_n}\frac{1}{m}S_{\varphi_k,\varphi_n}\theta_k m^{-(\theta_k+1)}dm \\
&= \sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\frac{1}{\varphi_k/\varphi_n}S_{\varphi_k,\varphi_n}\theta_k\int_1^\infty m^{-(\theta_k+2)}dm \\
&= \sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\frac{1}{\varphi_k/\varphi_n}S_{\varphi_k,\varphi_n}\theta_k\left[-\frac{1}{\theta_k+1}m^{-(\theta_k+1)}\right]_1^\infty \\
&= \sum_{k\in\{h,l\}}\frac{\theta_k}{\theta_k+1}\sum_{n\in\{h,l\}}\frac{1}{\varphi_k/\varphi_n}S_{\varphi_k,\varphi_n}.
\end{aligned}$$

The TFP misallocation statistic $\mathcal{M}$ is then

$$\begin{aligned}
\mathcal{M} &= \frac{e^{\sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\int\left[\ln\left(\frac{1}{\frac{\varphi_k}{\varphi_n}}\frac{1}{m}\right)S_{\varphi_k,\varphi_n}\theta_k m^{-(\theta_k+1)}\right]dm}}{\Lambda} \\
&= \frac{e^{\sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\int\left[\ln\left(\frac{1}{\frac{\varphi_k}{\varphi_n}}\right)S_{\varphi_k,\varphi_n}\theta_k m^{-(\theta_k+1)}+\ln\left(\frac{1}{m}\right)S_{\varphi_k,\varphi_n}\theta_k m^{-(\theta_k+1)}\right]dm}}{\Lambda} \\
&= \frac{e^{\sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\left[\ln\left(\frac{1}{\frac{\varphi_k}{\varphi_n}}\right)S_{\varphi_k,\varphi_n}+\int\ln\left(\frac{1}{m}\right)S_{\varphi_k,\varphi_n}\theta_k m^{-(\theta_k+1)}dm\right]}}{\Lambda} \\
&= \frac{e^{\sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\left[S_{\varphi_k,\varphi_n}\ln\left(\frac{1}{\frac{\varphi_k}{\varphi_n}}\right)-S_{\varphi_k,\varphi_n}\frac{1}{\theta_k}\right]}}{\Lambda} \\
&= \frac{e^{\sum_{k\in\{h,l\}}\sum_{n\in\{h,l\}}\left[S_{\varphi_k,\varphi_n}\left(\ln\left(\frac{1}{\frac{\varphi_k}{\varphi_n}}\right)-\frac{1}{\theta_k}\right)\right]}}{\Lambda}
\end{aligned}$$

where I have made use of

$$\int_1^\infty \ln\left(\frac{1}{m}\right)S_{\varphi_k,\varphi_n}\theta_k m^{-(\theta_k+1)}dm = \left[\frac{\theta_k\ln(m)+1}{\theta_k m^{\theta_k}}+C\right]_1^\infty = -\frac{1}{\theta_k}.$$

Alternatively note that this expression is equal to $-S_{\varphi_k,\varphi_n}E[\ln(\lambda^\Delta)|\varphi_k,\varphi_n]$. I have shown above that $\ln(\lambda^\Delta)$ conditional on the productivity gap is exponentially distributed with parameter θ_k. From the characteristics of an exponential distribution its expected value is $1/\theta_k$.

B.4 Moments of the markup distribution

Mean of markups (unweighted or sales weighted with Cobb-Douglas aggregator)

$$
\begin{aligned}
E[\mu] &= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \int_1^\infty \frac{\varphi_k}{\varphi_n} m S_{\varphi_k,\varphi_n} \theta_k m^{-(\theta_k+1)} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \frac{\varphi_k}{\varphi_n} S_{\varphi_k,\varphi_n} \theta_k \int_1^\infty m^{-\theta_k} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \frac{\varphi_k}{\varphi_n} S_{\varphi_k,\varphi_n} \theta_k \left[\frac{1}{1-\theta_k} m^{1-\theta_k}\right]_1^\infty \\
&= \sum_{k\in\{h,l\}} \frac{\theta_k}{\theta_k-1} \sum_{n\in\{h,l\}} \frac{\varphi_k}{\varphi_n} S_{\varphi_k,\varphi_n},
\end{aligned}
$$

where in the last equation it is assumed that $\theta > 1$, which is true if $\frac{\tau}{I} > \lambda - 1$. Otherwise the mean is ∞. Note that this is simply the mean of a Pareto distribution (once $\frac{\varphi_k}{\varphi_l} S_{\varphi_k,\varphi_l}$ is taken out of the integral). The geometric mean is computed from previously derived expressions:

$$E[\mu^{geo}] = e^{-\ln(\mathcal{M}\times\Lambda)}.$$

2nd moment of markups

$$
\begin{aligned}
E[\mu^2] &= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \int_1^\infty \left(\frac{\varphi_k}{\varphi_n} m\right)^2 S_{\varphi_k,\varphi_n} \theta_k m^{-(\theta_k+1)} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \left(\frac{\varphi_k}{\varphi_n}\right)^2 S_{\varphi_k,\varphi_n} \theta_k \int_1^\infty m^{1-\theta_k} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \left(\frac{\varphi_k}{\varphi_n}\right)^2 S_{\varphi_k,\varphi_n} \theta_k \left[\frac{1}{2-\theta_k} m^{2-\theta_k}\right]_1^\infty \\
&= \sum_{k\in\{h,l\}} \frac{\theta_k}{\theta_k-2} \sum_{n\in\{h,l\}} \left(\frac{\varphi_k}{\varphi_n}\right)^2 S_{\varphi_k,\varphi_n}.
\end{aligned}
$$

For the last equality to hold we need $\theta_k > 2$, otherwise the moment is ∞.

Variance of markups

$$
\begin{aligned}
Var(\mu) = E[\mu^2] - E[\mu]^2 = &\sum_{k\in\{h,l\}} \frac{\theta_k}{\theta_k-2} \sum_{n\in\{h,l\}} \left(\frac{\varphi_k}{\varphi_n}\right)^2 S_{\varphi_k,\varphi_n} \\
&- \left(\sum_{k\in\{h,l\}} \frac{\theta_k}{\theta_k-1} \sum_{n\in\{h,l\}} \frac{\varphi_k}{\varphi_n} S_{\varphi_k,\varphi_n}\right)^2
\end{aligned}
$$

Without any differences in productivity ($\varphi_h = \varphi_l$) and own-innovation rates $\theta_h = \theta_l$ the variance collapses to $\frac{\theta}{\theta-2} - \left(\frac{\theta}{\theta-1}\right)^2 = \frac{\theta}{(\theta-2)(\theta-1)^2}$, which is just the variance of the Pareto distribution μ collapses to.

Mean of log markups

$$
\begin{aligned}
E[\ln \mu] &= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \int_1^{\infty} \ln\left(\frac{\varphi_k}{\varphi_n} m\right) S_{\varphi_k,\varphi_n} \theta_k m^{-(\theta_k+1)} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \int_1^{\infty} \left(\ln\left(\frac{\varphi_k}{\varphi_n}\right) + \ln m\right) S_{\varphi_k,\varphi_n} \theta_k m^{-(\theta_k+1)} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \left(\ln\left(\frac{\varphi_k}{\varphi_n}\right) S_{\varphi_k,\varphi_n} + S_{\varphi_k,\varphi_n} \theta_k \int_1^{\infty} \ln(m) m^{-(\theta_k+1)} dm\right) \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \left(S_{\varphi_k,\varphi_n} \left(\ln\left(\frac{\varphi_k}{\varphi_n}\right) + \frac{1}{\theta_k}\right)\right)
\end{aligned}
$$

2nd moment of log markups

$$
\begin{aligned}
E[(\ln \mu)^2] &= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} \int_1^{\infty} \left[\ln\left(\frac{\varphi_k}{\varphi_n} m\right)\right]^2 S_{\varphi_k,\varphi_n} \theta_k m^{-(\theta_k+1)} dm \\
&= \sum_{k\in\{h,l\}} \sum_{n\in\{h,l\}} S_{\varphi_k,\varphi_n} \frac{\theta_k \ln\left(\frac{\varphi_k}{\varphi_n}\right)\left(\theta_k \ln\left(\frac{\varphi_k}{\varphi_n}\right) + 2\right) + 2}{\theta_k^2}
\end{aligned}
$$

The variance of log markups is then

$$
Var(\ln \mu) = E[(\ln \mu)^2] - E[(\ln \mu)]^2,
$$

which is computed using the above expressions. Without any differences in productivity ($\varphi_h = \varphi_l$) and own-innovation rates ($\theta_h = \theta_l$) the variance collapses to $1/\theta^2$, which is just the variance of the exponential distribution that $\ln \mu$ collapses to.

B.5 Deriving the steady-state growth rate of aggregate variables

The growth rate of aggregate variables is determined by the growth rate of Q_t.

$$
g = \frac{\dot{Q}_t}{Q_t} = \frac{\partial \ln(Q_t)}{\partial t}
$$

Quality of a product in a given product line increases through own-innovation, firm expansion or firm entry. For the growth rate of Q_t we have

$$
\begin{aligned}
\ln(Q_{t+\Delta}) &= \int_0^1 \ln(q_{t+\Delta,i}) di \\
&= \int_0^1 \Big[(\Delta S I^h + \Delta(1-S) I^l + \Delta S x^h + \Delta(1-S) x^l + \Delta z) \ln(q_{t,i}\lambda) \\
&\quad + (1 - \Delta S I^h - \Delta(1-S) I^l + \Delta S x^h + \Delta(1-S) x^l + \Delta z) \ln(q_{t,i})\Big] di \\
&= \int_0^1 \Big[(\Delta S I^h + \Delta(1-S) I^l + \Delta S x^h + \Delta(1-S) x^l + \Delta z) \ln(\lambda) + \ln(q_{t,i})\Big] di
\end{aligned}
$$

so that

$$\frac{\ln(Q_{t+\Delta})-\ln(Q_t)}{\Delta} = \Big(SI^h+(1-S)I^l+Sx^h+(1-S)x^l+z\Big)\ln(\lambda)$$
$$=\Big(SI^h+(1-S)I^l+\tau\Big)\ln(\lambda).$$

For $\Delta \to 0$, $g=\Big(SI^h+(1-S)I^l+\tau\Big)\ln(\lambda)$.

B.6 Markup dynamics

Firm markups are defined by $\mu_f = \frac{py_f}{wl_f} = \left(\frac{1}{n}\sum_{k=1}^{n}\mu_{kf}^{-1}\right)^{-1}$. Therefore

$$\ln\mu_f = -\ln\left(\frac{1}{n}\sum_{k=1}^{n}\mu_k^{-1}\right).$$

Rewrite the term in brackets (for a high productivity firm) as

$$\frac{1}{n}\sum_{k=1}^{n}\mu_k^{-1} = \frac{1}{n}\sum_{k=1}^{n}e^{-\ln\mu_k} = \frac{1}{n}\left(\sum_{i=1}^{n_i}e^{-\ln\frac{\varphi^h}{\varphi^l}-\ln\Delta_i\ln\lambda}+\sum_{j=1}^{n_j}e^{-\ln\Delta_j\ln\lambda}\right), \tag{34}$$

where i indexes the product lines where the high productivity firm faces a low productivity second best producer, j the lines where it faces a high productivity second best producer and $n_i+n_j=n$. A two-dimensional linear Taylor expansion around $\ln\lambda=0$ and $\ln\frac{\varphi^h}{\varphi^l}=0$ gives

$$\frac{1}{n}\left(\sum_{i=1}^{n_i}e^{-\ln\frac{\varphi^h}{\varphi^l}-\ln\Delta_i\ln\lambda}+\sum_{j=1}^{n_j}e^{-\ln\Delta_j\ln\lambda}\right)\approx 1-\left(\frac{1}{n}\sum_{k=1}^{n}\Delta_k\right)\ln\lambda-\frac{n_i}{n}\ln\left(\frac{\varphi^h}{\varphi^l}\right)$$

such that

$$E\Big[\ln\mu_f|\text{firm age}=a_f,\varphi^h\Big]\approx E\left[\frac{1}{n}\sum_{k=1}^{n}\Delta_k|\text{firm age}=a_f,\varphi^h\right]\ln\lambda+(1-S)\ln\left(\frac{\varphi^h}{\varphi^l}\right),$$

where I have used the fact that the expected share of the firm's product lines with a low productivity second best producer is equal to the aggregate share of product lines where the active producer is of the low productivity type. From Peters (2020) I know that

$$E\left[\frac{1}{n}\sum_{k=1}^{n}\Delta_k|\text{firm age}=a_f,\varphi^h\right]\ln\lambda=\Big(1+I^h\times E[a_P^h|a_f]\Big)\ln\lambda,$$

where $E[a_P^h|a_f]$ denotes the average product age of a high process efficiency firm conditional on firm age a_f and

$$E[a_P^h|a_f]=\frac{1}{x^h}\left(\frac{\frac{1}{\tau}\left(1-e^{-\tau a_f}\right)}{\frac{1}{x^h+\tau}\left(1-e^{-(x^h+\tau)a_f}\right)}-1\right)\Big(1-\phi^h(a_f)\Big)+a_f\phi^h(a_f)$$
$$\phi^h(a)=e^{-x^ha}\frac{1}{\gamma^h(a)}\ln\left(\frac{1}{1-\gamma^h(a)}\right)$$
$$\gamma^h(a)=\frac{x^h\left(1-e^{-(\tau-x^h)a}\right)}{\tau-x^he^{-(\tau-x^h)a}},$$

which gives the expression in the main text.

For a firm of the low process efficiency type the last term in eq. (34) reads

$$\frac{1}{n}\left(\sum_{i=1}^{n_i} e^{-\ln\Delta_i \ln\lambda} + \sum_{j=1}^{n_j} e^{-\ln\frac{\varphi^l}{\varphi^h} - \ln\Delta_j \ln\lambda}\right),$$

where i indexes the product lines where the low productivity producer faces a low productivity second best producer, j the lines where it faces a high productivity second best producer and $n_i + n_j = n$. Following the same steps as for a high productivity firm this time linearizing around $\ln\frac{\varphi^l}{\varphi^h} = 0$ (and $\ln\lambda = 0$) gives

$$E\left[\ln\mu_f|\text{firm age} = a_f, \varphi^l\right] \approx \left(1 + I^l \times E[a_P^l|a_f]\right)\ln\lambda + S\ln\left(\frac{\varphi^l}{\varphi^h}\right),$$

where again I have made use of the fact that the share of the firm's product lines with a high productivity second best producer is equal to the aggregate share of product lines where the active producer is of the high productivity type. $E[a_P^l|a_f]$ is exactly defined as $E[a_P^h|a_f]$ with x^h replaced by x^l in the above expressions.

C Testing model predictions

C.1 Determinants of revenue productivity

Fact 1: $TFPR$ increases independently in firm age and firm sales

That revenue productivity is increasing in either firm age or firm size has been documented in the literature, see, e.g., Peters (2020) presenting evidence of a positive relationship between inverse labor shares and firm age and Lentz and Mortensen (2008) showing a positive correlation between labor productivity and value added. The fact that revenue productivity relates positively to firm age and size does not necessarily imply that both drive it. The $TFPR$-size relationship is potentially driven by age: older firms are on average larger. Alternatively the $TFPR$-age relationship could be driven by size: larger firms are on average older.

To show that both age and size drive $TFPR$ I run a horse-race regression. I use lagged firm sales as the baseline measure for firm size. In the model I present in section 4 firm sales are proportional to the number of products a firm produces. I use lagged instead of contemporaneous sales to account for potential positive correlation between measurement errors in sales and value added entering the numerator of $TFPR$ that could lead to a spurious correlation between $TFPR$ and size. The main specification of the horse-race regression is

$$\ln TFPR_{fkt} = \beta_0 + \beta_1 Age_{fkt} + \beta_2 \ln Sales_{fkt-1} + \delta_{kt} + u_{fkt},$$

where $Sales_{fkt-1}$ denotes lagged firm sales and δ_{kt} captures industry $\times$ year fixed effects.

Table 10: Horse-race regression: revenue productivity, age and size

	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln TFPR_t$	$\ln TFPR_t$	$\ln \widehat{TFPR}_t$	$\ln TFPR_t$	$\ln TFPR_t$	$\ln TFPR_t$
Age_t	0.0090	0.0055	0.0065	0.0023	0.0132	0.0072
	(14.111)	(10.666)	(14.052)	(4.052)	(21.294)	(15.029)
$\ln Sales_t$	0.0607		0.0432			
	(26.554)		(18.923)			
$\ln Sales_{t-1}$		0.0371		0.0040		
		(15.563)		(1.234)		
$\ln Employment_t$					-0.0266	-0.0154
					(-15.632)	(-4.578)
Industry × year FE	✓	✓	✓		✓	
Firm FE				✓		✓
$Age_t < 10$	✓	✓	✓	✓	✓	✓
N	300,518	213,864	214,270	213,864	300,518	300,518

Notes: t statistics in parentheses. Clustered standard errors at the fixed effects level. $\ln \widehat{TFPR}$ denotes $\ln TFPR$ predicted by its lagged value (run for each firm). $TFPR$ is computed from balance sheet data, whereas employment is obtained from separate matched employer-employee data.

The regression output is shown in Table 10. Columns (1)-(3) show the $TFPR$, age and size relationship controlling for 5-digit industry × year fixed effects. Throughout the three columns the coefficients on age and size are positive and significant. Column (1) shows the relationship between $TFPR$, age and contemporaneous sales. The coefficient on age suggests that conditional on firm size, firms that are one year older than average firm age in an industry-year display 0.009 log points higher $TFPR$. At the same time, conditional on firm age, firms with one log point larger sales than industry-year average display 0.061 log points higher $TFPR$. The main specification is shown in column (2) with lagged sales as the size measure. Increasing age by one year increases $TFPR$ by 0.006 log points conditional on size and increasing size by one log point increases $TFPR$ by 0.037 log points conditional on age. The decline in the size coefficient from column (1) to (2) indicates that the size coefficient in column (1) is partly driven by the contemporaneous link between sales and value added (entering $TFPR$). However, both the age and size coefficient remain large and significant in column (2). To provide further robustness that the size coefficient is not driven by random fluctuations in $TFPR$ that correlate with sales I predict firm $\ln TFPR$ by its lagged value for each firm separately. Apart from breaking the link between contemporaneous value added and sales, predicting $TFPR$ by its lag smoothes the firm $TFPR$ profile. Column (3) shows the regression of predicted $TFPR$ on age and sales. The coefficients are comparable in size to column (2). Hence, in all specifications $TFPR$ is increasing in firm age and firm size independently across firms.

All specifications restrict to firms younger than ten years of age since the $TFPR$-age profile flattens out at age ten. See Table 11 in the Appendix for results of the horse-race regression including older firms. The size coefficient is close to the one reported for the main specification. In the same table

I further report results of separate $TFPR$-age and $TFPR$-size regressions and show that the above findings extend to using inverse labor shares instead of $TFPR$ as a measure of a firm's revenue productivity.

Fact 2: Controlling for firm fixed effects the $TFPR$-size relationship is small

The previous fact showed that firm age and size are important independent drivers of $TFPR$ across firms. I show next that an increase in size within the firm has a negligeable effect on $TFPR$.

Column (4) of Table 10 shows the main specification of the horse-race regression with firm instead of industry × year fixed effects. Whereas the age coefficient remains statistically significant, the size coefficient declines by an order of magnitude and turns insignificant. Within the firm a one log point increase in size is associated with a 0.004 log point increase in $TFPR$. The size effect is not only statistically, but also economically miniscule.

This result relates to Gamber (2021) who using Compustat data provides indirect evidence of markups increasing in sales. Regressing log variable input expenditures (e.g. wage bill) on log sales he finds coefficients smaller than one indicating that markups increase in sales. The rate at which variable input expenditures increase with sales is lower when controlling for firm than for industry × year fixed effects suggesting that markups increase faster in size within firms than across firms. This contrasts my findings. To relate my analysis to his I regress $\ln TFPR$ on lagged sales without controlling for age in columns (1)-(2) of Table 12 in the Appendix. The size effect on $TFPR$ within firms remains small relative to the across-firm size effect. I further show in columns (3)-(4) that also the effect of size on inverse labor shares declines by an order of magnitude when controlling for firm relative to industry × year fixed effects. To speak directly to his results I run his regression with my data regressing log wage bill on log sales and find a lower coefficient with industry × year fixed effects than with firm fixed effects, which is consistent with size having a larger effect on markups across firms than within. The results of this regression are shown in Table 13 columns (1)-(2), Appendix. One potential explanation for the difference in our findings is the coverage of firms. Using register data my analysis covers the universe of firms in Sweden. Gamber (2021) relies on Compustat data, which is restricted to publicly listed firms. In fact if I restrict my data to firms with sales exceeding 100 Mio. 2017 SEK (4.7% of all firms), I do find evidence that the size effect on markups is larger within firms than across (Table 13 columns (3)-(4)). That the size effect on the markup is larger across than within firms in the unrestricted sample, but not among the subset of very large firms is potentially driven by the fact that when looking at the complete cross-section of firms in an industry, variation in firm size is mainly driven by heterogeneity in firm fundamentals (e.g. process efficiency) that in return affects the markup. When narrowing in on large, successful firms in that industry those firms potentially look more alike in terms of firm fundamentals since, e.g., a high process efficiency level is required to reach that size. This would suggest that across firm differences in size have larger effects on markups in the complete cross section than when looking at a subset of either small or large firms. I repeat my regression of $TFPR$ on lagged sales with industry × year fixed effects for a subsample of small or large firms. Column (5) of Table 12 shows that for the subsample of large firms the size coefficient drops by half compared to the size coefficient for the complete cross section in column (1) and for the sample of small firms the coefficient is indistinguishable from zero, column (6). Hence, the large size coefficient in the complete cross-section is mainly driven by variation between large and small firms rather than by variation among small or large firms.

Fact 3: $TFPR$ increases faster in sales than employment

Columns (5) and (6) of Table 10 show results of the horse-race regression of $TFPR$ on age and employment. The measure for employment is obtained from a separate matched employer-employee data set ($RAMS$). While the age coefficient remains positive and significant the coefficient on employment is negative. Conditional on firm age firms with one log point larger employment relative to average employment in the industry-year display 0.027 log points lower $TFPR$. Within firms a one log point increase in employment is associated with 0.015 log points lower $TFPR$. I provide further robustness for the $TFPR$-employment relationship in Table 14 in the Appendix using both the wage bill to measure employment or using predicted $TFPR$. The sign and statistical significance of the coefficient on employment or the wage bill vary across specifications, however in all specifications is the coefficient on employment or the wage bill lower than the one reported for sales in the previous tables.

Stylized facts 1-3 suggest that first, firm age and size matter independently for markups, particularly for explaining across-firm differences. Second, conditional on firm fixed effects the increase in markups associated with an increase in firm size is small. That size differences across firms matter for explaining across firm differences in markups, but within-firm sales growth has a small effect on the markup suggests that across-firm differences in size are driven by firm fundamentals that affect the markup (e.g. differences in productivity), whereas within-firm sales growth arises from factors that are less related to markup growth. Third, markups increase less strongly in employment than in sales. These findings have clear predictions for modelling firm dynamics and in fact oppose some commonly used theories. First, markups varying with age and size contradict models where firms charge constant markups due to, e.g., a constant price elasticity of demand as in Melitz (2003). Second, stylized fact 1 highlights both firm age and sales as determinants of markups. Models that feature either a positive markup-age or markup-size relationship miss one dimension of markup heterogeneity. Further, the horse-race regression emphasizes age and size as *independent* determinants of markups. This opposes theories where a positive markup-age and markup-size relationship is driven by the same mechanism (e.g. through demand accumulation as in Foster, Haltiwanger and Syverson (2008)) or where the markup-age correlation simply arises through a positive age-size correlation. Holding size constant the horse-race regression predicts a positive relationship between markups and age. Conditioning on size holds the driver behind size differences constant (e.g. productivity), which implies that the markup-age relationship must arise from a separate mechanism. Third, fact 2 shows that increasing sales within the firm has little effect on the markup. This contrasts models where firms increase their markups as they grow in size due to, e.g., the price elasticity of demand decreasing in size either by construction or where this arises endogenously as in models of oligopolistic competition à la Atkeson and Burstein (2008). These models feature a positive within-firm markup-size correlation, whereas I find that this relationship is insignificant in the data. Fact 2 also relates to the story of demand accumulation. If accumulation of customer capital were behind the positive $TFPR$-size relationship, then within-firm size growth should be positively related to $TFPR$, which is rejected by fact 2.

C.2 Auxiliary regressions

Table 11: Revenue productivity, age and size

	(1) $\ln TFPR_t$	(2) $\ln TFPR_t$	(3) $\ln TFPR_t$	(4) $\ln TFPR_t$	(5) $\ln \frac{VA}{wL}$	(6) $\ln \frac{VA}{wL}$	(7) $\ln \frac{VA}{wL}$	(8) $\ln \frac{VA}{wL}$
Age_t	0.0127		0.0055	0.0026	0.0131		0.0044	0.0030
	(20.762)		(10.666)	(10.010)	(31.036)		(11.359)	(14.471)
$\ln Sales_{t-1}$		0.0390	0.0371	0.0337		0.0682	0.0668	0.0640
		(16.635)	(15.563)	(15.652)		(33.179)	(31.939)	(30.956)
Industry × year FE	✓	✓	✓	✓	✓	✓	✓	✓
$Age_t < 10$	✓	✓	✓		✓	✓	✓	
N	300,518	213,864	213,864	302,038	300,518	213,864	213,864	302,038

Notes: t statistics in parentheses. Clustered standard errors at the industry × year level. VA denotes value added and wL the wage bill.

Table 12: Revenue productivity and size, across and within firms

	(1) $\ln TFPR_t$	(2) $\ln TFPR_t$	(3) $\ln \frac{VA}{wL}$	(4) $\ln \frac{VA}{wL}$	(5) $\ln TFPR_t$	(6) $\ln TFPR_t$
$\ln Sales_{t-1}$	0.0390	0.0093	0.0682	0.0068	0.0204	0.0011
	(16.635)	(3.411)	(33.179)	(2.651)	(3.592)	(0.108)
Industry × year FE	✓		✓		✓	✓
Firm FE		✓		✓		
$Sales > 50$ Mio. 2017 SEK					✓	
$Sales < 5$ Mio. 2017 SEK						✓
N	213,864	213,864	213,864	213,864	23,998	29,273

Notes: t statistics in parentheses. Clustered standard errors at the fixed effects level. VA denotes value added and wL the wage bill.

Table 13: Wage bill and sales, across and within firms

	(1) $\ln wL_t$	(2) $\ln wL_t$	(3) $\ln wL_t$	(4) $\ln wL_t$
$\ln Sales_t$	0.6849	0.7082	0.7390	0.7271
	(115.879)	(52.659)	(55.572)	(30.175)
Industry × year FE	✓		✓	
Firm FE		✓		✓
$Sales > 100$ Mio. 2017 SEK			✓	✓
N	300,518	300,518	13,479	13,479

Notes: t statistics in parentheses. Clustered standard errors at the fixed effects level. wL denotes the wage bill.

Table 14: Horse-race regression with employment or wage bill

	(1) $\ln \widehat{TFPR}_t$	(2) $\ln \widehat{TFPR}_t$	(3) $\ln \widehat{TFPR}_t$	(4) $\ln \widehat{TFPR}_t$	(5) ln TFPR	(6) ln TFPR
Age_t	0.0081	0.0079	0.0031	0.0039	0.0080	0.0066
	(17.740)	(16.949)	(10.023)	(10.887)	(15.154)	(11.066)
$\ln Employment_t$	-0.0182		-0.0012			
	(-12.254)		(-0.576)			
$\ln wL_t$		0.0033		-0.0134		
		(1.956)		(-4.466)		
$\ln wL_{t-1}$					-0.0020	-0.0372
					(-1.197)	(-11.569)
Industry × year FE	✓	✓			✓	
Firm FE			✓	✓		✓
$Age_t < 10$	✓	✓	✓	✓	✓	✓
N	214,270	214,270	214,270	214,270	213,864	213,864

Notes: t statistics in parentheses. Clustered standard errors at the fixed effects level. $\ln \widehat{TFPR}$ denotes $\ln TFPR$ predicted by its lagged value (run for each firm).

D Decomposing the sources of economic growth

Incumbent creative destruction creates growth, yet lowers incentives for firms to engage in own-innovation. Is more creative destruction by incumbents beneficial or harmful for economic growth? The answer to this question depends on the relative contributions of creative destruction and own-innovation to aggregate growth. Similarly, firm entry generates growth, however, harms incumbent firms. Does more firm entry lead to more economic growth? The answer depends on the relative contributions by entrants and incumbents to aggregate growth.

In this section I decompose the aggregate growth rate into its contributions by incumbent own-innovation and expansion (differentiating by productivity type) and the contribution by firm entry. The expression for the growth rate in eq. (13), repeated below, naturally lends itself to such decomposition

$$g = \Bigg(\underbrace{\underbrace{SI^h + (1-S)I^l}_{\text{Incumbent own-innovation}} + \underbrace{Sx^h + (1-S)x^l}_{\text{Incumbent product expansion}}}_{\text{Incumbents}} + \underbrace{z}_{\text{Entrants}} \Bigg) \times \ln(\lambda).$$

The separate contributions to the growth rate are highlighted by the curly brackets. The exponents indicate the contributions by productivity type. Table 15 quantifies each component of the growth rate. Columns differentiate between own-innovation and creative destruction, whereas rows differentiate between incumbents (by productivity type) and entrants. Differentiating between innovation types, own-innovation makes up for the majority of aggregate growth (68.1%).

Creative destruction accounts for the remaining 31.9%. Between incumbents and entrants, incumbents account for almost all growth (95%). Entry generates relatively little growth (5%). Among incumbents own-innovation accounts for 71.7% of growth (creative destruction for 28.3%). Those numbers are similar to Peters (2020) who reports a share of aggregate growth accounted for by entrants of 4% and to Garcia-Macia, Hsieh and Klenow (2019) who find that among incumbents own-innovation accounts for 75%-80% of growth.[22] Differentiating further by productivity type, firms of the high productivity type account for 94% of economic growth generated by incumbents, while the low type accounts for 6%. High productivity type firms, making up 65% of the entering firms, but 94% of aggregate growth among incumbents hence overproportionally contribute to aggregate growth. Their larger contribution to aggregate growth is due to those firms endogenously innovating at faster rates than low productivity firms, which is separate from their contribution to aggregate productivity.

Table 15: Decomposing the sources of economic growth (initial BGP)

	Own-innovation	Creative destruction	
Incumbents φ^h	0.0146 (63.6%)	0.0059 (25.7%)	0.0205 (89.3%)
Incumbents φ^ℓ	0.0011 (4.6%)	0.0003 (1.1%)	0.0013 (5.7%)
Entrants		0.0012 (5%)	0.0012 (5%)
	0.0157 (68.1%)	0.0073 (31.9%)	0.0230 (100%)

Notes: table shows the contribution to growth by incumbents vs. entrants, by own-innovation vs. expansion and high (φ^h) vs. low (φ^h) productivity type for the initial balanced growth path (BGP). Percentages in brackets refer to the share of total growth (0.023).

The contributions to aggregate growth in eq. (13) derive from firms' innovation efforts that give rise to the revenue productivity (markup) and sales life cycle. The firm life cycle has been targeted explicitly in the model estimation. The reason why not more aggregate growth is generated through creative destruction is that this would impliy faster employment life cycle growth that in return is inconsistent with the life cycle growth targeted in the data. Similarly, for low productivity firms to contribute more to aggregate growth would required those firms to expand faster in employment, which again would be inconsistent with the data.

I perform the same decomposition for the new BGP following the fall in own-innovation costs (high type), the rise in entry costs and the rise in creative destruction costs. The results are shown in Table (16). In line with intuition, the share of aggregate growth accounted for by entrants declines along with the share accounted for by creative destruction. The contribution of own-innovation by the high productivity firms increases.

[22] Garcia-Macia, Hsieh and Klenow (2019) infer the contributions by studying employment fluctuations, while in Peters (2020) contributions to aggregate growth are deduced from firm life cycle growth as in this paper.

Table 16: Decomposing the sources of economic growth (new BGP)

	Own-innovation	Creative destruction	
Incumbents φ^h	0.0196 (70.7%)	0.0061 (21.8%)	0.0257 (92.5%)
Incumbents φ^ℓ	0.0009 (3.2%)	0.0002 (0.7%)	0.0011 (3.9%)
Entrants		0.0010 (3.7%)	0.0010 (3.7%)
	0.0205 (73.9%)	0.0073 (26.1%)	0.0278 (100%)

Notes: table shows the contribution to growth by incumbents vs. entrants, by own-innovation vs. expansion and high (φ^h) vs. low (φ^h) productivity type for the new balanced growth path. Percentages in brackets refer to the share of total growth (0.0278).

E Trading off static efficiency and growth

Utility from a consumption path where C_t grows at rate g is

$$\mathcal{U}\left(\{C_t\}_{t=0}^{\infty}\right) = \int_0^\infty e^{-\rho t} \ln C_t dt = \frac{1}{\rho} \ln C_0 + \frac{g}{\rho^2} = \mathcal{U}\left(C_0, g\right).$$

Utility of the consumption stream depends on the detrended consumption level C_0 and the growth rate. I evaluate the change in utility streams across two BGPs in permanent consumption-equivalent terms ξ as follows

$$\mathcal{U}\left((1+\xi)C_0^{old}, g^{old}\right) = \frac{\ln(1+\xi)}{\rho} + \mathcal{U}\left(C_0^{old}, g^{old}\right) = \mathcal{U}\left(C_0^{new}, g^{new}\right).$$

ξ measures the change in permanent consumption along the old BGP that equates utility of the old and new BGP. ξ then solves

$$\frac{\ln(1+\xi)}{\rho} = \frac{1}{\rho} \ln C_0^{new} + \frac{g^{new}}{\rho^2} - \frac{1}{\rho} \ln C_0^{old} - \frac{g^{old}}{\rho^2}.$$

Rearranging gives

$$\xi = exp\left(\frac{\ln\left(\frac{C_0^{new}}{C_0^{old}}\right)\rho + g^{new} - g^{old}}{\rho}\right) - 1.$$

The change in utility streams across BGPs depends on the discount rate, the relative detrended consumption levels and the difference in growth rates. To compute the relative detrended consumption levels I assume that both BGPs start off from the same average quality level. Differences in C_0 across BGPs then arise from differences in average productivity Φ, markup dispersion $\mathcal{M}$ and production labor L_P (all constant along each BGP)

$$\frac{C_0^{new}}{C_0^{old}} = \frac{\Phi^{new}}{\Phi^{old}} \times \frac{\mathcal{M}^{new}}{\mathcal{M}^{old}} \times \frac{L_P^{new}}{L_P^{old}}.$$

$\mathcal{M}$ and L_P are derived in the main text. The change in average productivity is given by

$$\frac{\Phi^{new}}{\Phi^{old}} = \left(\frac{\varphi^h}{\varphi^l}\right)^{S^{new}-S^{old}}.$$