Good Rents versus Bad Rents: R&D Misallocation and Growth

Philippe Aghion, Antonin Bergeaud

Timo Boppart, Peter J. Klenow, Huiyu Li*

Monday 27$^{\text{th}}$ May, 2024

Click here for the latest version

Abstract

Firm price-cost markups may reflect (a) bigger step sizes from quality innovations that confer significant knowledge spillovers onto other firms, and/or (b) higher process efficiency than competing firms. We write down an endogenous growth model in which, compared with the *laissez-faire* equilibrium, the social planner would generally like to reallocate research resources towards high markup firms in case (a) so as to capitalize on knowledge spillovers but not in case (b). We then exploit unit price variation across firms in French manufacturing to assess the relative strength of these two forces. Viewed through the lens of our model, the French data are consistent with significant variation in innovation step sizes, and hence gains from mitigating R&D misallocation. The policy implication is that, to reach the social optimum, French research subsidies should favor only those high markup firms with "good" rents.

*Aghion: Collège de France, INSEAD, and London School of Economics; Bergeaud: HEC Paris, CEP-LSE, and CEPR; Boppart: IIES, Stockholm University and University of St. Gallen; Klenow: Stanford University and the NBER; Li: Federal Reserve Bank of San Francisco. We thank Nicolas Cruzet, Maarten De Ridder and participants at the Conference in Memory of the work of Emmanuel Farhi and the NBER Summer Institute Macroeconomics and Productivity workshop and NBER EF&G Fall meeting for their helpful comments. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System, the Banque de France, or the Eurosystem. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 – Centre d'accès securisé aux données – CASD), and has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 786587). Philippe Aghion also acknowledges the support of EUR grant ANR-17-EURE-0001. Erin Crust provided excellent research assistance.

1. Introduction

A perennial debate weighs the static distortions created by price-cost markups against the dynamic incentives that such markups provide for private sector innovation. Classic examples include Dasgupta and Stiglitz (1980) and Tirole (1988). More recently, the debate has evolved to include firm heterogeneity in markups and innovation effort. Evidence for heterogeneous markups includes Edmond, Midrigan, and Xu (2015, 2021), De Loecker and Eeckhout (2018), Haltiwanger, Kulick, and Syverson (2018), De Loecker, Eeckhout, and Unger (2020), Baqaee and Farhi (2020), and Autor, Dorn, Katz, Patterson, and Van Reenen (2020).[1] Among reasons for heterogeneous research intensity are firm differences in process efficiency (Cavenaile, Celik, and Tian, 2021; Voronina, 2022), research productivity (De Ridder, 2019), or ability to implement innovations (Akcigit, Celik, and Greenwood, 2016; Ma, 2021).

Less attention has been paid to the possibility that the *source* of markups may differ across firms—with implications for the optimal allocation of research across such firms. Markups may differ because firms differ in the step size of their quality-improving innovations, such as in Klette and Kortum (2004) or Akcigit and Kerr (2018). Quality innovation can confer significant knowledge spillovers onto other firms, who can build on their innovations. Alternatively, markups may vary across firms because firms differ in their process efficiency (which arguably spills over to other firms less easily) or in their ability to circumvent regulations, obtain permits, or impose entry barriers on potential competitors (which bear no obvious knowledge externality).[2]

[1] A number of papers provide evidence for heterogeneous markups in the form of incomplete exchange rate pass-through. See Gopinath and Rigobon (2008), Gopinath, Itskhoki, and Rigobon (2010), Fitzgerald and Haller (2014), and Amiti, Itskhoki, and Konings (2019).

[2] One paper which documents heterogeneity in the source of markups is Akcigit, Baslandze, and Lotti (2018). Using firm-level evidence from Italy, these authors contrast small firms (whose rents stem from innovation) with market leaders (whose rents rely importantly on political connections). Only the former source of rents bestows knowledge spillovers on future innovators and generates growth externalities in their environment.

In this paper, we analyze the optimal allocation of R&D in an economy with multiple sources of markup heterogeneity across firms, and we contrast it with the market equilibrium under *laissez-faire*.[3] In the model we develop, firms differ both in their quality advantage over other firms and in their degree of proprietary process efficiency. Firms can improve quality through innovating, whereas the level of proprietary process efficiency of each firm is given once and for all. Both big quality steps and high process efficiency enable a firm to charge above-average markups, yet only big quality steps bestow knowledge spillovers on subsequent innovators (firms with better proprietary process efficiency do not in our setup).[4]

Our analysis sheds light on whether the allocation of research under *laissez-faire* is excessively or insufficiently tilted towards high-rent firms, depending upon whether the main source of markup heterogeneity across firms is quality steps versus proprietary process efficiency. The planner wishes to undo the static misallocation of production labor created by markup dispersion, but also endeavors to allocate research labor optimally. In particular, the planner wants to shift innovation effort toward big quality step firms and, as a byproduct, away from high process efficiency firms.

We use data on French manufacturing from 2012 to 2019 to infer the extent to which firms differ in their quality step sizes and process efficiency. According to our theory, big quality steps allow firms to raise their prices to charge high markups. High process efficiency, in contrast, allows firms to charge high markups by *not* passing through their lower marginal cost into

---

[3]To focus on the allocation of research across firms, we fix total research labor in this economy. That is, we set aside the question of whether the market devotes too little labor to research versus production.

[4]On the question of whether firms differ in their knowledge spillovers, the literature has drawn an important distinction between product innovation and process innovation. Product innovation appears easier for competing firms to reverse engineer and build upon. And process innovation appears easier for firms to keep secret from their competitors. This has historically tilted patenting toward products and away from processes—see the survey by Cohen and Levin (1989). Relatedly, Moser (2005, 2013) argues that the fraction of innovations that are patented varies over time and across industries depending on how much firms can keep their ideas secret if they do not patent them.

their prices. Thus, among firms with higher markups, high-price firms take big quality steps and low-price firms enjoy higher process efficiency.

We find that French manufacturing firms do indeed differ in their price-cost markups, as measured by the ratio of their revenues to their total costs. And firms differ in their unit prices conditional on these markups, suggesting heterogeneity in their quality steps versus process efficiency. When we infer the process efficiency of firms by dividing their markups by their price levels, we find that firm differences in process efficiency are quite persistent, consistent with our model. Furthermore, we find that cross-firm correlation of unit price with hours is positive, controlling for process efficiency while hours increases with process efficiency, controlling for average unit price.

Compared to the *laissez-faire* equilibrium, the planner's solution entails a roughly 41 percentage point shift in market share (the share of products operated) from low step size firms to firms with big step sizes. This speeds up the aggregate growth rate of the economy by 70 basis points, from about 2.3 to 3.0 percent per year. Compared to these dynamic gains, the static gains from eliminating the distorted allocation of production labor due to markup dispersion are second order.

This paper relates to several strands of literature: First, to the literature on competition, R&D, and growth such as Dasgupta and Stiglitz (1980), Aghion, Harris, and Vickers (1997), Aghion, Harris, Howitt, and Vickers (2001), and Aghion, Bloom, Blundell, Griffith, and Howitt (2005). Acemoglu and Akcigit (2012) use a step-by-step innovation model with leaders and followers to analyze the growth and welfare implications of various patent protection policies. They conclude that patents should provide stronger protection to firms with bigger technological leads over their rivals. There is only one source of markups in their model, however, namely quality differences. We contribute to this literature by considering multiple sources of markup heterogeneity at once, and by proposing an empirical method to infer the primary source of markup heterogeneity in an economy.

Second, our paper overlaps with the recent literature on whether market power and markup dispersion are inhibiting growth. Examples include Akcigit and Ates (2019), De Ridder (2019), Aghion, Bergeaud, Boppart, Klenow, and Li (2023), and Liu, Mian, and Sufi (2022). Our analysis is distinct in encompassing two sources of markup heterogeneity—process efficiency and quality steps—which allows us to characterize the socially optimal R&D allocation across firms with different markup sources and to contrast it with the *laissez-faire* equilibrium.

Third, two recent papers focus on R&D misallocation across fields and sectors. Acemoglu (2023), analyzes green versus dirty innovations, health care inventions emphasizing prevention versus ex post cures, etc. Liu and Ma (2021) compare the equilibrium and optimal R&D allocation across fields in a multi-sector environment involving dynamic cross-field knowledge spillovers. In contrast, our focus is R&D misallocation across firms within industries.

Recent papers focusing on R&D misallocation across firms include Babalievsky (2022), Lehr (2022), and Ayerst (2023). Babalievsky models market imperfections due to search frictions and spillovers. Lehr interprets dispersion in R&D intensity across firms as reflecting misallocation, and quantifies the extent to which growing R&D misallocation is undercutting U.S. growth. Ayerst contrasts private profitability (which leads to a high rate of patenting) with spillovers (proxied by patent citations). We focus instead on markup dispersion and its decomposition into quality (with spillovers) versus process efficiency (with no spillovers). We do not emphasize differences in patenting, patent citations, or R&D intensity across firms.

The rest of the paper proceeds as follows: Section 2. describes the facts that motivate our theory. Section 3. lays out our endogenous growth model of multiproduct firms with differing quality innovation step sizes and process efficiency levels. In Section 4. we calibrate the key model parameters based on the moments we document in the French data. Section 5. concludes.

2. Motivating facts

In this section we provide evidence that motivates the main features of our model. In particular, our model features: (i) heterogeneity in both process efficiency and quality across firms; (ii) product quality improvements over time, whereas process efficiency for each firm is given once and for all; (iii) quality innovations generate more spillovers onto other firms than process innovation. Feature (i) is motivated by our finding that a firm's size is positively associated with both its quantity per unit of inputs (its TFPQ) and its average unit price — each when controlling for the other variable. Feature (ii) is motivated by our documentation of meager within-firm growth in TFPQ. Finally, feature (iii) follows the extensive innovation literature and our own evidence that firms are much more likely to patent product innovations than process innovations. With these points in mind, we now detail the three facts in support of our main modeling assumptions.

2.1. Datasets

Our main source of data is the balance sheet and income statements of all firms that are subject to the standard corporate tax scheme in France (FARE).[5] From this data, we construct firm-level measures of nominal value added (VA), wage bill (W), and net tangible asset value (K). We merge FARE with the matched employer-employee data (DADS) to calculate a firm-level measure of hours worked (H), which we will use to address potential measurement errors in the FARE measure of labor inputs.

We focus on the manufacturing sector where product prices are available in the "Enquête Annuelle de Production" (EAP) dataset, which surveys

[5] We use the "consolidated" version of FARE which aggregate legal units (or "siren") into enterprise groups ("entreprises profilés"). This procedure mostly concerns the largest firms and allows a better consideration of their balance sheets. Indeed, the largest groups are often made up of numerous legal entities, each of them corresponding to an entry in the unconsolidated version of FARE. For this reason, we have to restrict the sample to 2012–2019.

manufacturing firms and covers all firms with more than 20 employees. For each firm, it splits sales firm sales PY into product-level sales py and units sold y.[6] Our cleaned dataset, with information on unit price (revenue/quantity), TFPR (revenue/inputs) and TFPQ (quantity/inputs), includes 13,155 firms with on average about 50,000 product-level observations per year. More details about the dataset and cleaning procedures are given in the Online Appendix A.

More specifically, we construct TFPQ for firm j in year t as its TFPR divided by its price:

$$\text{TFPQ}_{j,t} = \frac{\text{TFPR}_{j,t}}{P_{j,t}}.$$

Here, $\text{TFPR}_{j,t}$ is computed as the ratio of nominal value added to the geometric average of capital and labor inputs:

$$\text{TFPR}_{j,t} = \frac{VA_{j,t}}{K_{j,t}^{\alpha_{s(j,t),t}} W_{j,t}^{1-\alpha_{s(j,t),t}}}$$

where $s(j,t)$ is the industry of firm j in year t and $\alpha_{s,t}$ is the weighted average of the cost shares of capital across firms in industry s in year t:

$$\alpha_{s,t} = \frac{\sum_{s(j,t)=s} r_K K_{j,t}}{\sum_{s(j,t)=s} r_K K_{j,t} + W_{j,t}}.$$

We consider a breakdown of the manufacturing sector into 21 NACE-2-digit industries. We take $r_K = 0.1$ (10%) as a benchmark for the rental rate of capital.

To construct firm-level price $P_{j,t}$, we first construct unit prices for each product i sold by firm j in year t as $p_t(i,j) = p_t(i,j)y_t(i,j)/y_t(i,j)$.[7] We then calculate a firm-level price $P_{j,t}$ as the sales-weighted geometric average of

[6]Our product groups correspond to the classification used by the PRODCOM survey conducted by Eurostat. See Online Appendix A for more details. There are about 4,000 different product codes.

[7]Since prices are aggregated across products with potentially different units, we divide $p(i,j)$ by the weighted average of unit prices over all similar products (following De Ridder, Grassi, and Morzenti, 2022).

product unit prices:

$$P_{j,t} = \prod_{i=1}^{N_{j,t}} p_t(i,j)^{\omega_{i,j,t}}, \tag{1}$$

where $N_{j,t}$ is the number products sold by the firm and $\omega_{i,j,t} = p_t(i,j)y_t(i,j)/\sum_{i'=1}^{N_{j,t}} p_t(i',j)y_t(i',j)$ is the share of product i in firm j's sales in year t.

2.2. Facts

Fact 1. Firm size increases with firm price and firm TFPQ, each when controlling for the other variable We run a bivariate regression of hours worked on both TFPQ and price simultaneously. All variables are in logs and in deviations from industry-year fixed effects. We present the results in Table 1. The coefficient on TFPQ is 0.418 (standard error 0.056). Thus 1 higher log point of TFPQ goes along with about 42% higher hours worked across firms. Meanwhile, the coefficient on Price is 0.472 (standard error 0.061). So higher price firms are larger even controlling for their level of TFPQ. Note that as a proxy for size we have used hours worked from DADS ($H_{j,t}$) to avoid a by-construction correlation with the labor input used in the construction of TFPQ (the wage bill from FARE).

One might expect higher price firms to have lower sales. Perhaps the resolution to this puzzling pattern is that higher price firms sell higher quality goods, which are more costly to produce. Using this logic many studies use price as a proxy for product quality, e.g. Bils and Klenow (2004) and Hallak and Schott (2011).[8]

In Table 6 of the Appendix, we show that this pattern survives a whole set of robustness tests. In particular, Table 6 shows that the positive bivariate relationship is robust to controlling for age interacted with a sector fixed effect, to using five-digit sector fixed effects instead of two-digit fixed effects, and to

[8] Hottman, Redding, and Weinstein (2016) likewise find no lower costs for larger firms using AC Nielsen data on prices and market shares for consumer packaged goods manufacturers.

restricting attention on production workers when measuring total hours.

Table 1: Regression of firm size on firm TFPQ and price

	TFPQ	Price
Regression coefficients	0.418	0.472
Standard errors	(0.056)	(0.061)

Source: Underlying data is French manufacturing firm-year observations from 2012–2019. Price and TFPQ are in logs and are constructed from the FARE dataset, and the dependent variable is the log of hours worked from DADS. Control for industry-year fixed effects.

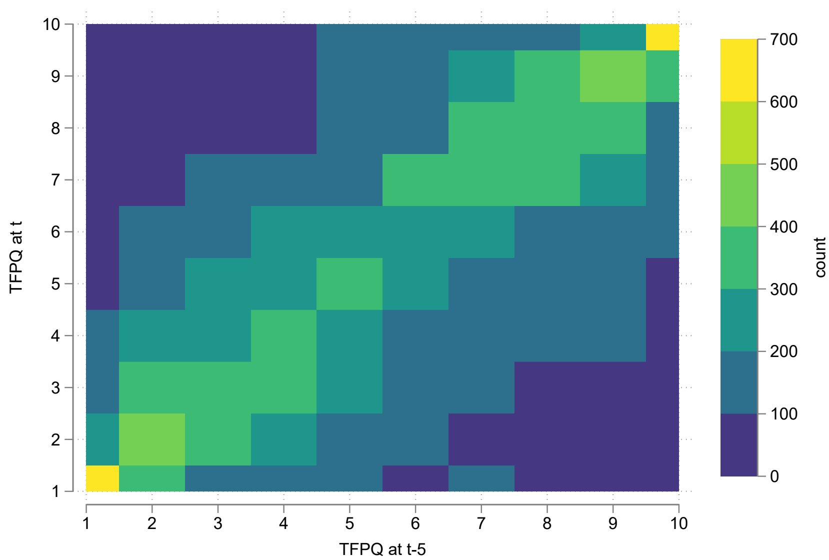
We take two lessons away from these facts. First, firms are importantly heterogeneous in their prices and process efficiencies. Second, these variables are not perfectly correlated with each other. So we would like a model with independent variation in quality and process efficiency, both of which affect prices and quantities.

Fact 2: Firm TFPQ exhibits little trend

Annual TFP growth averages 2.3% in French manufacturing over 1995–2019 (EU-KLEMS). This might stem in part from process efficiency growth. However, we find that within-firm TFPQ growth averages only 0.1% per year in French manufacturing from 2012–2019. Thus it appears that little growth comes from process innovation. As growth from new varieties is seldom captured in measured TFP growth (see for example Moulton (2018) and Aghion, Bergeaud, Boppart, Klenow, and Li (2019)), this leaves quality growth as the primary driver of measured TFP growth. This suggests that the priority is to model firm-level quality improvements.

Of course, we do find that TFPQ is heterogeneous across firms. So firms *do* seem to differ in their process efficiency. To see whether such differences are transitory or persistent, we ranked firms by their TFPQ and looked at the change in the rankings five years later. Figure 1 displays the transition matrix where the vertical axis is the firm's decile of TFPQ in year t and the horizontal axis is the firm's decile in year $t-5$. The color on the grid represents the number of firms. It shows that firms concentrate along the diagonal and firms tend to stay in the same decile of TFPQ over a 5-year window. Combined with the aforementioned evidence of low average growth in firm TFPQ, our data suggest that firm TFPQ is highly persistent. This motivates us to model firm process efficiency as given once and for all.

Figure 1: Firm TFPQ transition



Source: Underlying data is French manufacturing firm-year observations from 2012–2019. TFPQ is constructed from the FARE dataset. Ranking of TFPQ within industry-year.

Fact 3: Process innovation generates less spillovers than product innovation

Levin, Klevorick, Nelson, Winter, Gilbert, and Griliches (1987) provide the following explanation for why process innovations are less prone to generate knowledge spillovers than product innovation: *"The tendency to regard secrecy as more effective than process patents probably reflects the greater ease and desirability of maintaining secrecy about process technology. Maintaining secrecy about product innovations is thus likely to be both difficult and* *undesirable."*

This explanation finds support in the existing literature, for example, see Arundel and Kabla (1998), Cohen, Nelson, and Walsh (2000), Ornaghi (2006), Hall and Sena (2017), Banholzer, Behrens, Feuerriegel, Heinrich, Rammer, Schmoch, Seliger, and Wörter (2019), and Davison (2022). Among these papers, some suggest that process innovations are less often patented than product innovations, and we know that patents carry disclosure obligations which are meant to facilitate knowledge spillovers towards future potential innovators. Other papers look more directly at spillovers using citations, and argue that process innovations generate less follow up innovation than product innovations.

For our dataset on French manufacturing firms, we merged firm-level information on patenting with firm-level information on product versus process innovation. For each firm, we define "Patenting" as the event of obtaining a patent between 2014 and 2018 using the updated matching between PATSTAT and French firms from Bergeaud, Guillouzouic, Henry, and Malgouyres (2022). We merge this with the 2016 Community Innovation Survey (CIS) which surveys a random sample of firms on whether they have innovated on processes and products. For each firm, we set a dummy variable "Process" to 1 if the firm reports the introduction of a process innovation and a dummy variable "Product" to 1 if it reports the introduction of a product innovation.

We find that within the CIS sample, 9% of the firms reporting only process innovation (Process = 1, Product = 0) received a patent (Patenting = 1) compared to 28% of firms who reported only product innovation (Process = 0, Product = 1). By this measure, firms with product innovations are three times more likely to patent than those with process innovations. This is consistent with the argument in the literature that firms are more likely to keep their process innovations secret by not patenting them, and are more likely to disclose information on their product innovations through patenting.

The upshot from this patenting behavior is that spillovers may well be larger for product innovation than for process innovation. This interpretation is also consistent with Fact 2 wherein firm process efficiency differences appear to be fairly persistent over time. Of course, if process innovation contributes little to overall TFP growth, then that suggest there is less new knowledge to spill over from process innovation compared to product innovation.

3. Theory

We next lay out a theory that is consistent with the facts in the previous section. We pursue by characterizing the planner's solution and then by specifying a decentralized equilibrium. We next solve and compare the allocation of the planner to the decentralized outcome. We suppress time indices whenever it should not lead to confusion.

3.1. Setup

Household and preferences There is representative household with the following preferences over a final output good

$$\sum_{t=0}^{\infty} \beta^t \log(C_t). \tag{2}$$

Furthermore, the household can supply each period L units of production labor and Z units of R&D labor without generating disutility.

Final good production Final output, Y is a Cobb-Douglas bundle of a unit interval of intermediate goods which come at qualities $q(i)$s:

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

The final output good can be used for two purposes: it can be consumed, C, or—as will be explained below—used to cover production overhead cost, O.

Intermediate input production There is a "large" number of J firms which can produce the intermediate goods $i \in [0,1]$. Each firm produces at a line-specific quality level $q(i,j)$ and a firm-specific "process efficiency" $a(j)$. More specifically, a firm j can produce at their respective quality level $q(i,j)$ with constant labor productivity $a(j)$, i.e.,

$$y(i,j) = a(j) \cdot l(i,j), \tag{3}$$

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

The quality levels at which a firm produces change endogenously over time as a result of R&D activity. Each firm has access to a linear R&D technology with heterogeneous step sizes. That is, if $x \cdot \psi_z$ units of research labor are used by a particular firm j, this firm innovates in x randomly drawn lines. In a selected line the currently highest existing quality across firms is taken and increased by a factor $\gamma(j)$. The innovating firm j can then produce at this higher quality from the next period onward. The initial distribution of highest quality levels across firms is exogenously given. As the quality innovations are building on each other they therefore imply a positive spillovers on future innovators.

Overall the J firms in our model differ exogenously in two dimensions their level of process efficiency $a(j)$ and their innovation stepsize $\gamma(j)$. In the following we assume both dimension taking on two potential values (high H and low L in the case of productivity and big B and small S in the case of step size). We further assume $\gamma_S > a_H/a_L$.[9] Given the binary heterogeneity in both γ and a there are four type of firms: $k = \{HB, HS, LB, LS\}$. Here HS denotes a firms with a high process efficiency and a small innovation step size and so on.

[9]This will ensure that it is always optimal to produce by the highest quality firm in each line. This condition also ensures production in the highest quality firm (irrespective of its type) in the decentralized economy we study below.

On top of the linear production cost firms use resources on fixed production cost called "overhead". That is, in order to be active in a period a firm that has the highest quality level in $n(j)$ lines needs to spend $\psi_o\, n(j)^\eta Y$ units of final output on overhead. The curvature parameter η is greater than 1 such that overhead costs are convex in the number of lines.

In the following we assume that all firms of the same type k start out with the same number of line, $n_{k,0}$, in which they have the highest quality $q_0(i)$ across all firms. With the convex overhead cost schedule this ensures that it is optimal for the planner to always keep homogeneity within type and the planner's problem can be characterized in terms of four representative type of firms.[10] We assume that ϕ_k denotes the fraction of firms of type k.

Aggregates and resource constraints Aggregate resources used on overhead are then given by $O = \sum_{j=1}^{J} \psi_o\, n(j)^\eta Y$ and the economy's resource constraint reads

$$Y = C + O. \tag{4}$$

Next, we have the production and R&D labor resource constraints which are

$$Z = \psi_z \sum_{j=1}^{J} x(j) = \psi_z \sum_{k} J\phi_k x_k, \tag{5}$$

(where the second equality again exploits homogeneity within types) and

$$L = \sum_{j=1}^{J} \int_0^1 l(i,j)\, di. \tag{6}$$

[10]The same will also hold true in the decentralized equilibrium that we study. Along a balanced growth path, which we will study below, homogeneity within type is automatically fulfilled and this assumption does not put any additional restrictions.

Finally, we have an accounting equation that says that the total number of products of highest quality must sum to 1 across firms

$$\sum_k S_k = 1 \text{ where } S_k \equiv J\phi_k n_k. \tag{7}$$

3.2. Planner's problem

With homogeneity within the four type of firms $k = \{HB, HS, LB, LS\}$, the planner's problem can be characterized as follows:

$$\max_{\{C_t, Q_{t+1}, n_{k,t+1}, x_{k,t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log(C_t), \tag{8}$$

where $k = \{HB, HS, LB, LS\}$, subject to

$$C_t = Q_t \exp\left(\sum_k J\phi_k n_{k,t} \log(a_k)\right)\left(1 - \psi_o \sum_k J\phi_k (n_{k,t})^{\eta}\right) L, \tag{9}$$

$$Z = \psi_z \sum_k J\phi_k x_{k,t}, \tag{10}$$

$$Q_{t+1} = Q_t \exp\left(\sum_k J\phi_k x_{k,t} \log\left(\gamma_k\right)\right), \tag{11}$$

$$n_{k,t+1} = n_{k,t}\left(1 - \sum_{k'} J\phi_{k'} x_{k',t}\right) + x_{k,t}, \ \forall k, \tag{12}$$

and a given $Q_0 = \exp\left(\int_0^1 \log\left(q_0(i)\right) di\right)$, $n_{HB,0}$, $n_{LB,0}$, $n_{HS,0}$, and $n_{LS,0}$ and some non-negativity constraints

$$n_{k,t+1} \geq 0, \ x_{k,t} \geq 0, \ \forall k, t. \tag{13}$$

Equation (9) captures the resource constraint, i.e., that consumption equals output minus overhead. Here we already have exploited the fact that it is always optimal to set $l(i) = L$ due to the Cobb-Douglas technology.

Furthermore, we exploited that it is always optimal to produce by the highest quality firm in a given line.[11] Output can then be written as the product of the geometric mean quality, $Q_t \equiv \exp\left(\int_0^1 \log(q_t(i))\, di\right)$, the geometric average of process efficiency, $\exp\left(\sum_k J\phi_k n_{k,t} \log(a_k)\right)$, and L. The term $(1 - \psi_o \sum_k J\phi_k (n_{k,t})^\eta)$ captures output net of overhead. Equation (10) captures the constraint on researcher labor. Finally, (11) captures the law of motion of the average quality level and (12) gives the law of motion of the number of highest quality lines by each type of firm.

3.3. Decentralized economy

In the decentralized economy we assume competitive markets with the exception of intermediate good production. So final output production can be characterized as the behavior of a representative firm solving

$$\max_{\{y(i)\}_{i=0}^1} P \exp\left(\int_0^1 \log\left(q(i)y(i)\right) di\right) - \int_0^1 p(i)y(i)\, di. \tag{14}$$

In the following, we normalize the price of the final output $P = \exp\left(\int_0^1 \log\left(p(i)/q(i)\right) di\right)$ to one in all periods.

The representative household supplies inelastically L units of production labor and Z units of research labor to the labor market and solves

$$\max_{\{C_t, A_{t+1}\}_{t=1}^\infty} \sum_{t=0}^\infty \beta^t \log(C_t), \tag{15}$$

subject to $A_{t+1} = A_t(1 + r_t) + w_t L + w_{z,t} Z - C_t$, $\forall t$ and a standard no-Ponzi game condition. Here A denotes wealth, r the interest rate and w and w_z the wage rates of production and research labor, respectively.

Intermediate input production The J intermediate input producers own patents to produce at particular qualities in given lines. The distribution of

[11] This follows from the restriction $\gamma_S > a_H/a_L$ we made above.

qualities across lines increases endogenously due to innovations. In each line the different firms that produce at different quality levels then compete à la Bertrand. We solve at this point already for the equilibrium pricing decision under Bertrand competition at intermediate input level. We will then arrive at an expression for period profits of a firm and this allows us to directly focus on the dynamic firm problem in isolation.

As we assumed $\gamma_S > a_H/a_L$, the firms that has the highest quality patent in a line is the "leading" firm, i.e., the firm with the lowest quality-adjusted marginal cost. Due to the Cobb-Douglas structure, under Bertrand competition, it is always optimal for this leading firm j in a given line i to set its quality-adjusted price equal to the quality-adjusted marginal cost of the second-best quality producer j', i.e., $\frac{p(i,j(i),j'(i))}{q(i,j(i))} = \frac{w}{q(i,j'(i))\cdot a(j')}$. This price setting implies the following markup factor over marginal cost charged by producer $j(i)$ in a line i

$$\mu(i, j(i), j'(i)) = \gamma_j \frac{a(j(i))}{a(j'(i))}. \tag{16}$$

The markup is equal to the step size (which depends on the identity of the producing firm) times the ratio of process efficiency of the producing firm relative to the process efficiency of the second-best firm. So the process efficiency of the second-best firm, which is either high or low, does influence markups (and therefore profits from a given line). Due to the Cobb-Douglas structure of final output production, sales in each product line are given by Y and independent of the quality level and prices. Hence, operating profits of a producing firm in a given line (before overhead cost) are given by $Y(1 - 1/\mu(i, j(i), j'(i)))$.

Total period profit of firm j then depend on the number of lines in which the firm has the highest quality patent, $n(j)$, and the share of these lines in which they face a high productivity second-best firm $h(j)$. These two variables $n(j)$ and $h(j)$ are the two individual state variables in the dynamic firm problem. So total profits after overhead expressed relative to output Y are given by

$$\pi\left(j,n,h\right)=nh\left(1-\frac{a_H}{\gamma_j a_j}\right)+n(1-h)\left(1-\frac{a_L}{\gamma_j a_j}\right)-\psi_o n^{\eta}. \tag{17}$$

When we again assume homogeneity within type, (17) can be expressed as four type-specific profit functions $\pi_k(n,h)$ for $k = \{HB, HS, LB, LS\}$ that depend on the individual state variables $n(j)$ and $h(j)$. The profit functions are type-specific as they depend on a_j and γ_j of the producing firm j (see (17)). The dynamic firm problem can then be expressed as follows:

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

subject to

$$n_{t+1}=n_t(1-Z/\psi_z)+x_t,\ \forall t, \tag{18}$$

$$h_{t+1}n_{t+1}=h_t n_t(1-Z/\psi_z)+S_t x_t,\ \forall t, \tag{19}$$

for a given n_0 and h_0 and some non-negativity constraints $x_t \geq 0$, $n_{t+1} \geq 0$. That is, each firm is choosing the R&D activity such to maximize the net present value of firm profits. The constraints (18) give the dynamics of number of lines in which the firm produces. There the firms take the aggregate rate of creative destruction Z/ψ_z as exogenously given. The second set of constraints (19) gives the change in the share of line in which the firm faces a high process efficiency second-best firm. In this process the firm take into account that the newly innovated lines are drawn from the total pool of line in which a fraction S is currently served by high process efficiency firms.

Market clearing and aggregates We have the labor markets clearing conditions (5) and (6) and the asset market clearing condition

$$\sum_k J\phi_k V_k = A.$$

Finally, we have the accounting equation (7) and an accounting equation giving the aggregate share of lines served by high process efficiency firms

$$S_t = S_{HB,t} + S_{HS,t} = J\left(\phi_{HB} n_{HB,t} + \phi_{HS} n_{HS,t}\right). \tag{20}$$

Equilibrium definition A decentralized equilibrium is then defined as a sequence of quantities and prices that jointly solve the final producer problem, the intermediate producer problems, the household problem and is in line with market clearing and all the aggregate constraints.

3.4. Definition of a BGP

We define a balanced growth path (BGP) in the standard way, i.e., as a path along which all quantities grow at constant rates. In the following we focus on an interior such BGP in which growth is strictly positive and all the type of firms are active. Under some parameter restriction such an interior BGP exists and is unique (both in the planner's problem as well as for the decentralized equilibrium). We next characterize the BGP in the planner's solution and then move on to the decentralized equilibrium and compare the two. We will denote variables along the BGP by upper bar, i.e., $\bar{r}$ will denote the interest rate along the decentralized BGP. Along a BGP, in both the planner's solution as well as the decentralized equilibrium, the distribution of the number of lines across firms will be stationary and output, consumption, the geometric mean of quality, $Q_t = \exp\left(\int_0^1 \log\left(q_t(i)\right) di\right)$ and total resources used for overhead grow at some endogenous rate $\bar{g}$.

3.5. Welfare along the BGP

As consumption grows along the BGP at a constant endogenous rate $\bar{g}$ welfare in (2) can be rewritten as

$$\frac{1}{1-\beta}\left(\log(\bar{C}_0)+\frac{\beta}{1-\beta}\log(1+\bar{g})\right), \tag{21}$$

where $\bar{C}_0 \equiv C_t(1+\bar{g})^{-t}$ is the detrended consumption level along the BGP. That is, welfare can be written as a weighted sum of the logarithm of the consumption *level* and *growth*, where the relative weight put on the logarithm of the gross growth rate is $\frac{\beta}{1-\beta}$. As output and innovations are both produced from distinct type of labor which are in fixed exogenous supply (L and Z), discrepancies along the BGP in the consumption level $\bar{C}_0$ and the growth rate $\bar{g}$ between the decentralized equilibrium and planner's solution solely arise from differences in the allocation of the fixed L and Z resources across the heterogeneous firms. In this sense our model focuses entirely on misallocation, both statically and in terms of R&D resources and shuts down potential distortions on the amount of resources devoted to R&D.

In the decentralized equilibrium as well as the planner's solution the detrended consumption level along the BGP can be written as the following product

$$\bar{C}_0 = (1-\bar{o})\cdot Q_0\cdot\Phi\cdot\mathcal{M}\cdot L, \tag{22}$$

where $\bar{o} \equiv O/Y = \psi_o \sum_k J\phi_k(\bar{n}_k)^{\eta}$ is the fraction of output used for overhead, $Q_0 = \exp\left(\int_0^1 \log\left(q_0(i)\right)di\right)$ is the initial geometric mean quality level, $\Phi = \exp\left(\sum_k \bar{S}_k \log\left(a_k\right)\right) = a_S\Delta^{\bar{S}}$ is the geometric average of process efficiency, and $\mathcal{M}$ captures potential misallocation of labor across lines. The last term $\mathcal{M}$ is equal to one for the social planner, whereas it is smaller than one due to markup dispersion across lines in the decentralized equilibrium. In the decentralized equilibrium this allocative efficiency term is the ratio of geometric relative to the arithmetic average of the inverse markups across lines, or formally

$$\mathcal{M} = \frac{\exp\left(\int_0^1 \log \frac{1}{\mu(i,j(i),j'(i))}\, di\right)}{\int_0^1 \frac{1}{\mu(i,j(i),j'(i))}\, di} \leq 1. \tag{23}$$

The terms $\bar{o}$ and the aggregate process efficiency level Φ are both functions of the distribution of lines provided by the different type of firms $\bar{S}_k$. Overhead resources $\bar{o}$ are minimized if all firms are of equal size $n_j = 1/J, \forall j$ irrespective of their type (which implies $\bar{S}_k = \phi_k, \forall k$). In contrast, aggregate process efficiency Φ is maximized if the high process efficient firms serve the whole market, i.e., $\bar{S} = 1$. As the decentralized equilibrium and the planner's solution will give rise to a different market share distribution along the BGP (and there is markup dispersion in the decentralized equilibrium) the detrended consumption level $\bar{C}_0$ will differ between the two solution concepts.

Similarly to the level, also the growth rate can be expressed as a function of the market share distribution along the BGP. We have

$$1 + \bar{g} = \exp\left(\frac{Z}{\psi_z}\left[\sum_k J\phi_k \bar{n}_k \log(\gamma_k)\right]\right) = (\bar{\gamma})^{\frac{Z}{\psi_z}}, \tag{24}$$

where $\bar{\gamma} \equiv \Pi_k \gamma_k^{\bar{S}_k}$, and $\bar{S}_k \equiv J\phi_k \bar{n}_k$. Here, $\bar{\gamma}$ is the geometric mean of the step size γ_k weighted by the share of lines $\bar{S}_k$ served by type k firms. This equation again holds true for both the decentralized equilibrium as well as the planner's solution. Growth rate is higher when more products are produced by firms with big step size. The intuition for this result is the following: as total research labor is fixed and all firms have the same linear R&D technology the rate of creative destruction is equal to $\frac{Z}{\psi_z}$. Now along a balanced growth path—in order for the firm size distribution to be stationary—all the types of firms have to innovate at a rate that is proportional to their market share $\bar{S}_k$. The growth rate is then simply given by the market share-weighted geometric average of the step size raised to the rate of creative destruction $\frac{Z}{\psi_z}$.

This shows how the main ingredients into welfare—the detrended consumption level plus the growth rate—are both a function of the market share distribution across firms. In order to produce, the quality improvements have to be developed in house and this creates an interesting trade-off. The planner weighs off the level and growth effect as highlighted in (21). In contrast, the market share distribution along the decentralized BGP is determined by the relative profitability (capability to charge markups) across firms which will generally lead to inefficiencies. We will next characterized the optimal market share distribution chosen by the planner along the BGP and then contrast it with the decentralized equilibrium.

3.6. Characterizing the BGP of the planner's solution

How does the planner determine the market shares $\bar{S}_k = J\phi_k\bar{n}_k$? As shown in the Online Appendix B, the optimal differences between any two $\bar{n}_k$ and $\bar{n}_{k'}$ satisfy

$$\bar{n}_k^{\eta-1} - \bar{n}_{k'}^{\eta-1} = \frac{1-\bar{o}}{\psi_o\eta}\left[\log\left(\frac{a_k}{a_{k'}}\right) + \left(1 + \frac{Z/\psi_z}{1/\beta - 1}\right)\log\left(\frac{\gamma_k}{\gamma_{k'}}\right)\right], \tag{25}$$

where $\bar{o} = O/Y$ is again the output share of total overhead costs along the BGP. This result says that the planner chooses a higher long-run market share for firms with higher process efficiency a and for firms with bigger step size γ. The reason for this is that increasing the number of products in high a firms increases the aggregate level of process efficiency Φ in production, whereas increasing the number of products in big γ firms has on top of a similar level effect also a positive effect on the long-run growth rate $\bar{g}$.[12] As a consequence, the relative weight the planner puts on differences in step sizes versus process efficiency is increasing in the consumer's patience (β) or in the amount of available research labor (Z/ψ_z). The differences in the optimal long-run number of products across firm are moderated by the scalar in front of the overhead cost curve. If the overhead cost curve shifts down (as ψ_o decreases)

[12] Recall that choosing a high long-run market share for one firms also implies that this firm uses a larger share of the R&D labor.

the optimal long-run differences in firm size are magnified.

The level of $\bar{n}_k$ is determined by combining (25) with the accounting equation (7) on the aggregate number of products. For example, if $\eta = 2$ the overhead cost is quadratic, and the planner's first-order condition (26) becomes

$$\bar{n}_k - \bar{n}_{k'} = \frac{1-\bar{o}}{2\psi_o}\left[\log\left(\frac{a_k}{a_{k'}}\right) + \left(1 + \frac{Z/\psi_z}{1/\beta - 1}\right)\log\left(\frac{\gamma_k}{\gamma_{k'}}\right)\right], \tag{26}$$

and the optimal share of products produced by type-k firms is given by

$$\bar{S}_k = \phi_k\left(1 + \nu_k \frac{\sqrt{1 + (\frac{J}{\psi_o} - 1)\sum_k \phi_k \nu_k^2} - 1}{\sum_k \phi_k \nu_k^2}\right). \tag{27}$$

Here,

$$\nu_k \equiv \log\left(\frac{a_k}{\prod_{k'} a_{k'}^{\phi_{k'}}}\right) + \left(1 + \frac{Z/\psi_z}{1/\beta - 1}\right)\log\left(\frac{\gamma_k}{\prod_{k'} \gamma_{k'}^{\phi_{k'}}}\right)$$

summarizes the technology level of type-k firms relative to the geometric mean across firms. Since the differences in γ have additional implications for the long-run growth rate, the term in front of $\log(\gamma_k)$ exceeds 1 and the planner places more weight on step size differences.[13]

As we noted in the setup of the planner's problem, the planner finds it optimal to allocate the same amount of labor to each product line. Hence, $\bar{S}_k$ is also the share of labor allocate to type-k firms or $L_k/L = \bar{S}_k$. Furthermore, since the rate of innovation in each firm $\bar{x}_k/\bar{n}_k$ is equal to the aggregate rate of

[13] We assume that parameters are such that there is an interior BGP. For $\eta = 2$ this is guaranteed as long as the following condition holds

$$0 < \frac{\sqrt{1 + (\frac{J}{\psi_o} - 1)\sum_k \phi_k \nu_k^2} - 1}{\frac{J}{2\psi_o}\sum_k \phi_k \nu_k^2} < 1. \tag{28}$$

This is a necessary and sufficient condition for the solution given by (27) to generate $\bar{o} \in (0,1)$. When this condition is satisfied, we can show that the share of products $\bar{S}_k$ approaches the share of firms ϕ_k when the overhead cost schedule shifts up (higher ψ_o). The convex overhead cost is the reason why the planner does not allocate all products to a particular type of firm with the highest combination of step size and process efficiency.

creative destruction and all firms have the same R&D efficiency, $\bar{S}_k$ is also the share of R&D labor allocated to type-k firms.

3.7. Characterizing the BGP of the decentralized equilibrium

The first-order condition for firms in the decentralized equilibrium implies

$$\bar{n}_k^{\eta-1} - \bar{n}_{k'}^{\eta-1} = \left(\frac{a_L}{\gamma_{k'} a_{k'}} - \frac{a_L}{\gamma_k a_k}\right) \frac{\bar{S}\Delta + 1 - \bar{S}}{\psi_o \eta}, \quad \bar{S} \equiv \bar{S}_{HS} + \bar{S}_{HB}. \tag{29}$$

Unlike the planner's condition, the firms do not distinguish between step size and process efficiency. For example, if $\eta = 2$, we show in Online Appendix B that the number of products produced by a type-k firm along an interior BGP satisfies

$$\bar{n}_k = \frac{1}{J} + \frac{\bar{S}(\Delta - 1) + 1}{\gamma_S \psi_o 2} \omega_k, \tag{30}$$

where $\bar{S}$ is the share of products produced by firms with high process efficiency, Δ is the process efficiency gap a_H/a_L and

$$\omega_k \equiv \left(\sum_{k'} \phi_{k'} \frac{\gamma_S a_L}{\gamma_{k'} a_{k'}}\right) - \frac{\gamma_S a_L}{\gamma_k a_k}$$

is the markup advantage of a type-k firm vis-à-vis an average firm. Firms with higher markups, either due to bigger step size or higher process efficiency, have higher than average share of products. This results in the product shares in the decentralized equilibrium

$$\bar{S}_k = \phi_k \left(1 + \omega_k \frac{(\Delta - 1)(\phi_{HB} + \phi_{HS}) + 1}{\gamma_S \frac{2\psi_o}{J} - (\Delta - 1)(\phi_{HB}\omega_{HB} + \phi_{HS}\omega_{HS})}\right). \tag{31}$$

The decentralized product shares in equation (31) differ from the expression for the planner's product shares. In contrast to the planner's solution, the relative market shares are just pinned down by the relative markups and are independent of the relative weight β consumers put on growth and level effects. As an increase in the minimum step size γ_S shifts—for a given Δ and Γ—the markup distribution to the right, γ_S has an influence on the decentralized market shares along the BGP. Such an effect is not in the planner's solution, where the optimal market shares only depends on relative step size and relative process efficiency Γ and Δ.

Furthermore, the relative labor share of two firms j and k facing the same share of high process efficiency competitors $\bar{S}$ is given by

$$\frac{\lambda_j(\bar{S})}{\lambda_k(\bar{S})} = \frac{\gamma_k a_k}{\gamma_j a_j}, \tag{32}$$

as all firms are facing along the BGP the same share of high process efficiency competitors $\bar{S}$. Since production wages are the same across firm and $\bar{S}_k$ is equal to the sales share of type-k firms, the relative employment of type k firms given by

$$\frac{L_k}{L_{k'}} = \frac{\lambda_k(\bar{S})\bar{S}_k}{\lambda_{k'}(\bar{S})\bar{S}_{k'}} = \frac{\gamma_{k'} a_{k'}}{\gamma_k a_k}\frac{\bar{S}_k}{\bar{S}_{k'}}. \tag{33}$$

This expression says that firms with high markups (high $\gamma_k a_k$) have employment shares that are lower than their sales and product shares. This is a contrast to the planner's solution, where employment and market share coincide for any given firm.

On top of this firm level difference, there is also heterogeneity in labor allocated to the different production lines *within* a firm along the competitive BGP. This is because the markup across production lines within a firm differs by a factor Δ depending on whether the second-best firm is of high or low process efficiency. The amount of labor a firm devotes to line i where it faces a high efficiency competitor relative to the amount of labor it devotes to line i' where it faces a low efficiency competitor is therefore given by

$$l(i,j) = \Delta\, l(i',j). \tag{34}$$

This gives rise to an additional static efficiency loss as the planner would equalize the amount of production labor across all lines within a firm. The heteorogeneity in labor allocation across lines lowers the level of detrended consumption relative to the planner's allocation and shows up as $\mathcal{M} < 1$ in our welfare decomposition (21). The allocative efficiency term becomes along the BGP

$$\mathcal{M} = \frac{\Delta^{\bar{S}}}{\Delta\bar{S} + 1 - \bar{S}} \frac{\exp\left(-\sum_k \bar{S}_k \log(a_k\gamma_k)\right)}{\sum_k \bar{S}_k \frac{1}{a_k\gamma_k}} < 1.$$

Finally, as in the planner's problem, the share of research labor allocated to type-k firms is the same as the share of products produced by type-k firms because the arrival rate of new products matches the aggregate rate of creative destruction $\bar{z} = Z/\psi_z$ for all firms.

3.8. Planner's solution versus the decentralized equilibrium

As we summarized above, in general, the planner's allocation is different from the decentralized allocation. First, conditional on a given number of product lines in which a firm is the highest quality producer, the planner and decentralized equilibrium differ in the allocation of employment across product lines. Equation (33) implies that the planner always want more employment in the firms with higher $\gamma_k a_k$ than the decentralized equilibrium, conditional on the product shares. Both step size and process efficiency gaps generate markup and labor share heterogeneity across firms of different $\gamma_k a_k$. On top of that the type of second-best firms will additionally generate variations in production labor across lines within firms in the decentralized equilibrium. Along the BGP each firm has a share $\bar{S}$ of lines in which they face a high type second-best in in such lines production labor is higher compared to the remaining $1 - \bar{S}$ lines by a factor of Δ. In contrast, the planner wants to allocate the same amount of labor to each line. Second, the decentralized equilibrium in general will deviate from the planner's product shares leading to static and/or dynamic misallocation. We will use three extreme cases to illustrate this.

First, let us consider the case where step sizes are the same across firms $\gamma_k = \gamma$. In this case, there is no dynamic misallocation in the sense that the growth rate is $1 + \bar{g} = \gamma^{\frac{Z}{\psi_z}}$ in both the decentralized equilibrium and the planner's problem. However, the planner's share of products allocated to the high process efficiency firms ($\bar{S}_H^P$) can differ from the share in the decentralized equilibrium ($\bar{S}_H^D$). So this is an extreme case where efficiency boils down to only static efficiency. We show in Online Appendix B that when $\eta = 2$

$$\frac{\bar{S}_H^P - \phi_H}{\bar{S}_H^D - \phi_H} = \frac{(\frac{J}{\psi_o} - 1)\,(\log\Delta)\,\frac{\frac{2\psi_o}{J}\frac{\gamma\Delta}{\Delta-1} - (\Delta-1)\phi_H(1-\phi_H)}{(\Delta-1)\phi_H+1}}{\sqrt{1 + (\frac{J}{\psi_o} - 1)\,(\log\Delta)^2\,(1-\phi_H)\phi_H} + 1}, \tag{35}$$

which implies that $\bar{S}_H^P - \bar{S}_H^D$ increases with the common step size γ. In the decentralized equilibrium, the gap in profit share between the high and low types shrinks with γ and $\bar{S}_H^D$ approaches ϕ_H as γ goes to infinity. The step size does not affect planner's $\bar{S}_H^P$.

Hence, when firms have the same step size but different process efficiency levels, the long-run growth rate is the same in the planner's solution and the decentralized equilibrium but the level of consumption can be lower in the decentralized equilibrium due to static misallocation (differences in employment and product shares). These differences result in differences in the share of overhead costs $\bar{o}$, aggregate process efficiency $\Phi = \Delta^{\bar{S}_H}$ and allocative efficiency (markup dispersion) $\mathcal{M}$ in decomposition (21).

Another polar case is where all firms have the same process efficiency $a_k = a$. This is equivalent to setting $\Delta = 1$. Define ϕ_B as the share of firms with γ_B and $\bar{S}_B$ the share of products they produce. For both the planner and decentralized equilibrium, the geometric-mean of the step sizes is given by

$$\bar{\gamma} = \Gamma^{\bar{S}_B}\gamma_S.$$

As a consequence, the growth rate increases with the share of products produced by firms with the big step size $\bar{S}_B$.

We show in Online Appendix B for the simple case with $\eta = 2$ that

$$\frac{\bar{S}_B^P - \phi_B}{\bar{S}_B^D - \phi_B} = \frac{(\frac{J}{\psi_o} - 1)\left(\frac{1/\beta - 1 + Z/\psi_z}{1/\beta - 1}\log\Gamma\right)\frac{\gamma_B}{\Gamma - 1}\frac{2\psi_o}{J}}{\sqrt{1 + (\frac{J}{\psi_o} - 1)\left(\frac{1/\beta - 1 + Z/\psi_z}{1/\beta - 1}\log\Gamma\right)^2 (1 - \phi_B)\phi_B} + 1}, \tag{36}$$

which implies that $\bar{S}_B^P - \bar{S}_B^D$ increases in γ_B holding fixed Γ. In the decentralized equilibrium, profit shares approaches 1 as γ_B and γ_S increase. Hence, the gap in profit share between the big and small types shrinks and $\bar{S}_B^D$ approach ϕ_B when γ_B increases while Γ stays constants. However, the planner only cares about Γ and $\bar{S}_B^P$ does not change when γ_B increases while Γ stays constant. Therefore, when firms differ in step sizes, growth rates in the decentralized equilibrium will in general deviate from the growth rate in the planner's problem.

A final special case is the one were the product $a_k \cdot \gamma_k$ is the same across all firms. This is the case when $\Delta = \Gamma$ and process efficiency and the step size are perfectly negatively correlated, i.e., $\phi_{HB} = \phi_{LS} = 0$. In this case, as markups are equalized across firms the number of lines per firm along the decentralized BGP is equalized or $\bar{n}_{LB}^D = \bar{n}_{HS}^D = 1/J$ (see (30) and note that ω_k is equal to zero for both groups k). Interestingly, this is an allocation that minimizes overhead cost $\bar{o}$. However, the decentralized equilibrium does not take into account the dynamic positive externality generated by the big step size firms. The planner would indeed along the BGP choose a larger market share of the big step size firms LB and a lower market share for the high process efficiency firms HS. Formally, we have

$$(\bar{n}_{LB}^P)^{\eta-1} - (\bar{n}_{HS}^P)^{\eta-1} = \frac{1 - \bar{o}}{\psi_o \eta}\frac{Z/\psi_z}{1/\beta - 1}\log(\Gamma) > 0. \tag{37}$$

Hence, the planner would indeed sacrifice some static process efficiency and increase the overhead cost $\bar{o}$ to instead exploit the larger growth potential of the big step size firms. The extent to which this is done depends on the shape of the overhead cost function (ψ_o and η), the growth potential of the economy determined by the available R&D labor Z/ψ_z and the discount rate β.

3.9. Decentralized equilibrium with R&D tax policies

Before we move onto calibrating the model, we first extend the model to allow for R&D subsidies so that we can incorporate the effect of these programs on our target data moments and calculate policy counterfactuals.

In France, businesses face a flat income tax rate $\tau = 33\%$. They receive subsidy $\underline{\tau}_{RD} = 30\%$ for R&D expenditure up to 100 million Euros and a $\overline{\tau}_{RD} = 5\%$ for R&D expenditure in excess of the cutoff. Incorporating these elements, the post tax period profit relative to Y for the intermediate producers becomes

$$\begin{aligned} &(1-\tau)\left(\pi_k(n,h) - \frac{\psi_{z,k}w_z}{Y}x\right) && \text{post income tax} \\ + \;& \underline{\tau}_{RD}\min\left\{\frac{\psi_{z,k}w_z}{Y}x,\, \underline{RD}\right\} && \text{subsidy for R\&D below threshold} \\ + \;& \overline{\tau}_{RD}\max\left\{\frac{\psi_{z,k}w_z}{Y}x - \underline{RD},\, 0\right\} && \text{subsidy for R\&D in excess threshold} \end{aligned}$$

where $\underline{RD}$ is the ratio of the 100 million Euros cutoff relative to total output.

The R&D subsidy adds an additional determinant of firms size. Let $\tau_{RD}(x,\psi_{z,k})$ denote the marginal subsidy rate for a firm that spend R&D to draw x lines. The marginal rate can be expressed as

$$\tau_{RD}(x,\psi_{z,k}) = \begin{cases} \underline{\tau}_{RD} & \text{if } \frac{\psi_{z,k}w_z}{Y}x < \underline{RD} \\ \overline{\tau}_{RD} & \text{if } \frac{\psi_{z,k}w_z}{Y}x > \underline{RD} \end{cases} \tag{38}$$

The first-order conditions for the firms become

$$\bar{n}_k^{\eta-1} - \bar{n}_{k'}^{\eta-1} = \frac{\bar{S}\Delta + 1 - \bar{S}}{\psi_o \eta \gamma_S}\left(\frac{\gamma_S a_L}{\gamma_{k'} a_{k'}} - \frac{\gamma_S a_L}{\gamma_k a_k}\right) \quad \text{markup difference}$$

$$+ \frac{\psi_z w_z}{Y}\left(1/\beta - 1 + Z/\psi_z\right)\frac{\tau_{RD}(Z/\psi_z \bar{n}_k) - \tau_{RD}(Z/\psi_z \bar{n}_{k'})}{1-\tau} \quad \text{R\&D subsidy}$$

If the marginal R&D subsidy is the same for all firms in the equilibrium, the subsidy does not affect the equilibrium firm size distribution. It only raises the equilibrium w_z and pretax R&D expenditure. The post tax R&D expenditure is the same. However, since $\overline{\tau}_{RD}(x) < \underline{\tau}_{RD}(x)$, the marginal subsidy is lower for firms with high enough $\bar{n}_k$. Hence the R&D subsidy makes larger firms smaller compared to an economy with a flat subsidy.

Then net tax revenue T is the difference between revenue from the business income tax and outlay for R&D subsidies. We assume the net tax revenue is rebated lump sum to the household.

4. Calibration

In this section, we calibrate 9 parameters—$\psi_o, Z/\psi_z, \beta, \gamma_S, \gamma_B, \Delta, \{\phi_{k,k'}\}$—in the BGP of the decentralized equilibrium. We set η the curvature parameter of the overhead cost to 2. We will also address measurement error that may contaminate measures of firm-level markups, prices and productivity by calibrating the variance of classical measurement error in firm level price and TFPR. We choose structural and measurement error parameters to fit 1) variation in markups (TFPR), price and productivity (TFPQ) across firms, 2) the elasticity of size (total firm hours) with respect to firm TFPQ and price, 3) aggregate markup, productivity growth rate, interest rate, and R&D share of output.[14]

We will first provide a heuristic description of how the aforementioned

[14]We only need 12 moments for 13 parameters because we have the restriction $\sum_k \phi_k = 1$.

moments discipline the parameters in the model. Then we describe how we address measurement errors in the data. After describing the overall strategy of the calibration, we will summarize how we construct the target moments from firm-level and product-level data. Finally, we will show the calibration results and use the calibrated parameter values to compare allocations under the decentralized equilibrium with the planner's allocation.

4.1. Intuition for calibration

Here, we will provide some intuition for how certain moments are informative of particular parameters.[15] First, recall that in the decentralized equilibrium, the price of good i is given by

$$p(i, j(i), j'(i)) = \frac{w \cdot \gamma(j(i))}{a(j'(i))},$$

where $j(i)$ and $j'(i)$ index respectively the producing and the second-best firm. We can calculate firm j's price index as the sales-weighted average of the prices of all products produced by the firm. Along a BGP firms innovate upon a randomly drawn line from a stationary pool of types of producing firms. Hence, along the BGP, the share of products produced by type k firms ($\bar{S}_k$) is the same as the share of products of any firm where the second-best producer is of type k. Using these two properties of the model, the "average" price level for a firm j is

$$p_{j,t} = \bar{w}_0(1+\bar{g})^t \cdot \gamma_j \sum_{k'} \frac{\bar{S}_{k'}}{a_{k'}} \propto \gamma_j. \tag{39}$$

Equation (39) says that along the BGP, cross-firm price variation is entirely driven by cross-firm differences in the step size of innovations γ_j. The heterogeneity in process efficiency does not explain any of the differences in the price level across firms as such productivity differences just affect the

[15] It should be noted however that the parameters are calibrated jointly. The model is nonlinear in parameters and hence the effect of a parameter on outcomes in general depends on the value of other parameters.

markups but leave prices unaffected due to Bertrand competition. This gives us the following prediction about the dispersion of prices *across* firms

$$Var_j(\log(p_j)) = Var_j(\log(\gamma_j)) = (\phi_{HB} + \phi_{LB})(\phi_{HS} + \phi_{LS})(\log \Gamma)^2, \quad (40)$$

where Var_j is the variance operator using the distribution of firm types. Given the share of firms with big step sizes ($\phi_{HB} + \phi_{LB}$), the dispersion of prices across firms is increasing in the gap in step sizes Γ. Given Γ, dispersion increases as the share of firms with big step sizes approaches half.

Next, we define firm-level TFPR as firm-level revenue over firm-level costs and firm-level TFPQ as TFPR divided by the firm-level price index we described earlier. Since we do not have physical capital in our model, TFPR is proportional to firm markup which is the inverse of firm-level labor share. Recall that the labor share of a firm j along the BGP is given by

$$\lambda_j = \frac{\bar{S}a_H + (1 - \bar{S})a_L}{\gamma_j a_j} \quad (41)$$

where $\bar{S} = \bar{S}_{HB} + \bar{S}_{HS}$ is the sales-share of high process efficiency firms. Hence, cross-firm differences in TFPR are driven both by differences in the step sizes of innovation and by process efficiency heterogeneity i.e.

$$\text{TFPR}_j \propto \gamma_j a_j. \quad (42)$$

As a consequence, TFPQ of a firm—defined as its TFPR divided by firm price level—is proportional to its process efficiency:

$$\text{TFPQ}_j \equiv \frac{\text{TFPR}_j}{p_j} \propto a_j. \quad (43)$$

Therefore, dispersion in firm-level TFPQ is given by

$$Var_j(\log(\text{TFPQ}_j)) = Var_j(\log(a_j)) = (\phi_{HB} + \phi_{HS})(\phi_{LB} + \phi_{LS})(\log \Delta)^2 \quad (44)$$

and is informative of the gap in process efficiency Δ. All else equal, TFPQ dispersion increases with Δ and as the share of high process efficiency firms $(\phi_{HB} + \phi_{HS})$ approaches half.

Furthermore, given the dispersion in TFPQ and prices, dispersion in TFPR is informative about the covariance of step sizes and process efficiency across firms as we have

$$\begin{aligned} & \frac{Var_j(\log(\text{TFPR}_j)) - Var_j(\log(\gamma_j)) - Var_j(\log(a_j))}{2} \\ = \;& Cov_j(\log(\gamma_j), \log(a_j)) \\ = \;& (\phi_{HB}\phi_{LS} - \phi_{HS}\phi_{LB})(\log \Gamma)(\log \Delta). \end{aligned} \tag{45}$$

Given the gap in step sizes and process efficiency, the covariance increases with $\phi_{HB}\phi_{LS} - \phi_{HS}\phi_{LB}$. For example, the covariance is negative if the distribution of types has more weight on big step size and low process efficiency firms than big step size and high process efficiency firms $(\phi_{HS}\phi_{LB} > \phi_{HB}\phi_{LS})$.

The aggregate cost-weighted markup is informative about the step size γ_L given relative step sizes Γ and process efficiency Δ. Aggregate markup along the BGP satisfies

$$\frac{\bar{Y}}{wL} = \frac{1}{\bar{S}\Delta + 1 - \bar{S}} \frac{\gamma_S}{\sum_k \bar{S}_k \frac{\gamma_S}{\gamma_k} \frac{a_L}{a_k}}. \tag{46}$$

Hence an increase in γ_S shifts—for given relative step sizes Γ and relative process efficiency Δ—the entire markup distribution to the right.

Finally, as shown in (31), the sales share distributions across firms depend on the underlying distribution of types. Hence, we use the dispersion and skewness (median relative to mean) of the sales share distribution together with the covariance of TFPQ and firm level prices as three moments for calibrating the value of ϕ_{HB}, ϕ_{HS} and ϕ_{LB}. The value of ϕ_{LS} is determined by one minus the sum of these three values.

4.2. Accounting for measurement errors

The previous section described our general strategy for disciplining the parameters in our model. Before carrying out the calibration, we need to also lay out a strategy for dealing with measurement errors in the data as studies found large measurement errors in firm-level measured TFPR and prices (see for example Bils, Klenow, and Ruane, 2021). Here we lay out a strategy that addresses classical multiplicative measurement errors commonly used in the literature.

Let $\hat{v}$ denote the measured value of variable v. Suppose measured price, TFPR and TFPQ for a firm j is assumed to be related to the true price, TFPR and TFPQ as follows

$$\ln \widehat{p}_j = \ln p_j + \epsilon_j^p \tag{47}$$

$$\ln \widehat{\text{TFPR}}_j = \ln \text{TFPR}_j + \epsilon_j^{\text{TFPR}} \tag{48}$$

$$\ln \widehat{\text{TFPQ}}_j \equiv \ln \frac{\widehat{\text{TFPR}}_j}{\widehat{p}_j} = \ln \text{TFPQ}_j + \epsilon_j^{\text{TFPR}} - \epsilon_j^p \tag{49}$$

where ϵ_j^{TFPR} and ϵ_j^p are independent of each other, p and TFPR. Note that we construct measured TFPQ by dividing measured TFPR by measured price.

The dispersion in measured price, TFPR and TFPQ across firms are given by

$$\begin{aligned}
Var_j(\ln \widehat{p}_j) &= Var_j(\ln \gamma_j) + Var_j(\epsilon_j^p) \\
Var_j(\ln \widehat{\text{TFPR}}_j) &= Var_j(\ln \gamma_j + \ln a_j) + Var_j(\epsilon_j^{\text{TFPR}}) \\
Var_j(\ln \widehat{\text{TFPQ}}_j) &= Var_j(\ln a_j) + Var_j(\epsilon_j^{\text{TFPR}}) + Var_j(\epsilon_j^p)
\end{aligned}$$

while

$$\frac{Var_j(\ln \widehat{\text{TFPR}}_j) - Var_j(\ln \widehat{p}_j) - Var_j(\ln \widehat{\text{TFPQ}}_j)}{2} = Cov_j(\ln \gamma_j, \ln a_j) - Var_j(\epsilon_j^p).$$

Therefore, the dispersion in measured price and TFPQ across firms overstates the true dispersion in step sizes and process efficiency. Also, the gap between the dispersion in measured TFPR and the dispersion in measured prices and TFPQ understates the true covariance between step sizes and process efficiency when there are large measurement errors in prices. Hence, we need to know the degree of measurement errors $Var_j(\epsilon_j^p)$ and $Var_j(\epsilon_j^{\mathrm{TFPR}})$ to correctly infer the parameters in the model.

How do we gauge the extent of measurement errors? For the firms in our sample, we have a measure of labor input that is from a separate source and is not used to construct TFPR.[16] We construct a firm's labor input $\widehat{l_j}$ using this measure. Suppose $\widehat{l_j}$ deviates from the true labor input of a firm l_j by a classical multiplicative error

$$\ln \widehat{l_j} \quad = \quad \ln l_j + \epsilon_j^l, \tag{50}$$

where measurement error ϵ_j^l is independent of the measurement errors in prices and TFPR as well as step sizes and process efficiency. Given parameters, the model implies a relationship between a firm's employment, a firm's price and TFPQ through the relationship between employment share, step sizes and process efficiency. Measurement errors attenuate this relationship towards zero. Therefore, we can project measured employment from the independent source onto measured prices and TFPQ to generate additional moments to pin down the degree of measurement errors in prices and TFPQ.

4.3. Baseline calibration

Table 2 displays the data moments that we use to calibrate the parameters of the model. We will briefly describe how we calculate the moments and refer the reader to Online Appendix D for details. First, we target the slope coefficients obtained when we regress the log firm hours share within an industry-year on log of firm-level prices and log of firm-level TFPQ, controlling

[16]We construct this measure using the matched employer-employee data (DADS) which we aggregate at the firm level, while TFPR is constructed using only information from the firm's balance sheet using FARE. See the Online Appendix A for more detail.

for industry-year fixed effects. We calculate the target for dispersion in measured price, TFPQ and TFPR by first calculating firm-level price, TFPQ and TFPR as described in the previous section and then calculating their dispersion across firms within industry-year groups. We calculate aggregate dispersion in each year by averaging industry dispersion with industry cost-share weights.

We obtain the remaining target values from external sources. For average markups, we use the estimate from De Ridder et al. (2022) over 2009–2019 for France manufacturing. We take average annual rate of manufacturing productivity growth and R&D expenditure share of output from EU-KLEMS over 1995–2019 and the real interest rate for U.S. from 1996 to 2016 as estimated by Farhi and Gourio (2018).

Table 3 display the calibrated parameter values where we give equal weights to all moments. The regression coefficients of labor input on firm price and TFPQ in Table 2 are informative about the correlation between step sizes and process efficiency across firms. Since in the model price and TFPQ are proportional to process efficiency and step size respectively, these coefficients ask for parameters such that on average firm size increases with firm step size but declines with firm process efficiency. However, all else equal, firm employment increases with both process efficiency and step size. Hence, the model needs a negative correlation between step size and process efficiency (or $\phi_{HB}\phi_{LS} < \phi_{LB}\phi_{HS}$) to fit the empirical relationship between size, price and TFPQ. It also needs the step size gap Γ to be bigger than the process efficiency gap Δ so that the largest firms have big step size. We find that we need $\Gamma = 1.33$ and $\Delta = 1.13$.

Given Γ, the dispersion in price increases when $\phi_{HB} + \phi_{LB}$ approaches half. Therefore, the dispersion of price in the data as well as the coefficients jointly pins down ϕ_{HB} to 0.06 and ϕ_{LB} to 0.15.

The calibrated level of γ_S is 1.31, much lower than the target on aggregate markup level because γ_B is much bigger than γ_S and a large fraction of firms have big step size.

We need the discount factor β to be 0.97 to fit the interest rate target and R&D costs ψ_z/Z to be 16.8 to fit the productivity growth rate. The model picks ψ_o/J = 0.09 to fit the R&D share of output. All else equal, a higher ψ_o/J implies a lower marginal return to R&D and hence lower value of R&D labor w_z. Finally, as a by-product of our calibration, we find that measurement error in prices is about 97% of the observed dispersion in firm-level prices while measurement error in TFPR is 83% of the dispersion in measured firm-level TFPR.

Table 2: Baseline calibration

Targets	Data	Model
1. Dispersion in firm-level prices, $Var_j(\log \hat{p})$	0.436	0.450
2. Dispersion in firm-level TFPQ, $Var_j(\log \widehat{\text{TFPQ}})$	0.588	0.518
3. Dispersion in firm-level TFPR, $Var_j(\log \widehat{\text{TFPR}})$	0.088	0.094
4. Semi-elasticity of firm employment share wrt firm price, control TFPQ	0.472	0.467
6. Semi-elasticity of firm employment share wrt firm TFPQ, control price	0.418	0.415
7. Variance of residual from regressing firm labor input on firm price	0.419	0.443
8. Variance of residual from regressing firm labor input on firm TFPR	0.084	0.080
9. Aggregate price-cost markup ratio	1.5	1.44
10. Productivity growth rate (ppt year)	2.3	2.3
11. Interest rate (ppt/year)	5.2	5.2
12. R&D share of output (percent)	10.6	10.6

Sources: 1 to 8: authors' calculations from DADS, EAP and FARE, French manufacturing, 2012–2019. 9: De Ridder, Grassi, and Morzenti (2022), sales-weighted harmonic markup (equal to cost-weighted markup in the model) for France manufacturing 2009–2019. 10 and 12: EU-KLEMS, French manufacturing, TFP growth in labor-augmenting form, 1995–2019. 11: Farhi and Gourio (2018), U.S. all economy, 1996–2016.

4.4. Welfare decomposition

We next evaluate the decomposition in equation (21) at the calibrated parameters to compare welfare in the decentralized equilibrium with the

Table 3: Baseline calibrated parameters

Parameters	Parameter definitions	Values
ϕ_{HB}	Share of firms with high process efficiency and big step size	0.06
ϕ_{HS}	Share of firms with high process efficiency and small step size	0.32
ϕ_{LB}	Share of firms with low process efficiency and big step size	0.15
ϕ_{LS}	Share of firms with low process efficiency and small step size	0.47
γ_S	Small step size	1.31
$\Gamma \equiv \gamma_B/\gamma_S$	Step size gap	1.33
$\Delta \equiv \varphi_H/\varphi_L$	Process efficiency gap	1.13
ψ_o/J	Scale of overhead cost	0.09
β	Discount factor	0.97
ψ_Z/Z	R&D cost relative to R&D labor	16.8
	Implied share of measurement error in prices	97%
	Implied share of measurement error in TFPR	83%
	Implied share of measurement error in hours	2.5%

social optimum. Table 4 displays the components in equation (21) and Table 5 compares the allocation of products by firm types. The first row of Table 4 displays the distance of the decentralized economy to the first best in consumption-equivalent terms. Namely, we calculate the percent increase ξ in the consumption level $\bar{C}_0$ in the decentralized equilibrium such that welfare is the same as the planner's allocation or formally

$$\log(1+\xi) = \log\left(\frac{\bar{C}_0^P}{\bar{C}_0^D}\right) + \frac{\beta}{1-\beta}\log\left(\frac{1+\bar{g}^P}{1+\bar{g}^D}\right).$$

Table 4: Welfare comparison

Welfare gain in consumption-equivalent terms ξ	9.4%
Difference in the growth rate (ppt) $\bar{g}^P - \bar{g}^D$	0.71
Relative consumption level $\bar{C}_0^P / \bar{C}_0^D$	0.82
Relative consumption share $(1-\bar{o}^P)/(1-\bar{o}^D)$	0.81
Relative process efficiency $\bar{\Phi}^P / \bar{\Phi}^D$	0.99
Relative allocative efficiency $\bar{\mathcal{M}}^P / \bar{\mathcal{M}}^D$	1.01

The planner chooses a higher growth rate (3.0 vs. 2.3 ppt) than in the decentralized equilibrium as it is optimal to allocate more products (78.8 vs. 37.9 percent) to big step size firms as shown in Table 5. On the other hand, the consumption level $\bar{C}_0$ is lower in the planner's solution as the planner sacrifices some resources on overhead and allocates smaller market shares to high process efficiency firms in order to exploit the growth potential of big step size firms. The higher overhead cost share of the planner arises because it is optimal to have more products produced by big step size firms. In Table 5, the planner increases the number of products produced by each big step size firm (3.7 vs. 1.8) even though they are already larger on average than high process efficiency firms in the decentralized equilibrium (1.8 vs. 1.3). Finally, the planner has slightly higher allocative efficiency. This term is small because $\gamma_k a_k$ does not vary much due to the negative correlation between step sizes and process efficiency and as a consequence there is not so much markup dispersion. We need to take a stand on $\bar{Q}_0$ to calculate the overall welfare gain. To keep our gains conservative, we assume $\bar{Q}_0$ is the same. Overall, the welfare gain is 9.4% in consumption-equivalent terms.

Table 5: Product share (in percent) and firm size, planner vs. decentralized

	S_{HB}	S_{HS}	S_{LB}	S_{LS}	high proc. eff. share	big step size share
Decentralized	12.5	34.7	25.4	27.4	47.2	37.9
Planner	24.6	17.8	54.2	3.4	42.4	78.8

	n_{HB}	n_{HS}	n_{LB}	n_{LS}	high proc. eff. avg n	big step size avg n
Decentralized	2.1	1.1	1.7	0.6	1.3	1.8
Planner	4.1	0.6	3.6	0.1	1.2	3.7

Notes: The first panel displays the share of products produced by firms of each type. "High proc. eff share" = $\bar{S}_{HB} + \bar{S}_{HS}$ and "big step size share" = $\bar{S}_{HB} + \bar{S}_{LB}$. The $\bar{n}_k$'s are the number of products produced by each firm type relative to the mean $1/J$.

5. Conclusion

In this paper, we characterized the optimal research allocation in an economy where markup heterogeneity may be due to both differences in the step size of quality innovations and to differences in process efficiency across firms. To the extent that, unlike process efficiency, quality innovations confer knowledge spillovers onto other firms, we find the social planner will tilt innovation effort toward big quality step firms to enhance knowledge spillovers and thereby growth. At the same time, the planner will seek to undo the static misallocation of production labor created by markup dispersion.

We used data on French manufacturing firms from 2012 to 2019 to calibrate our model, and inferred a significant amount of independent variation in both the step size of innovations and process efficiency across firms. As a result, the planner can achieve sizable growth and welfare gains from reallocating research from high process efficiency firms to big step size firms. A corollary is that research subsidies should not favor all large firms in the context of the French economy—only those with big step sizes.[17]

There are additional sources of firm heterogeneity that we did not model that can also affect R&D misallocation. Examples include firm differences in research efficiency such as Luttmer (2011), about which the planner may have less information as in Akcigit, Hanley, and Stantcheva (2022). This could generate size differences that are unrelated to markup differences.

Our paper features a representative consumer, but could be extended to feature heterogeneity of firm ownership. In this way our theory and empirics could be extended to connect to a growing literature on income and wealth inequality such as Boar and Midrigan (2022), Aghion, Akcigit, Bergeaud, Blundell, and Hémous (2019), Song, Price, Guvenen, Bloom, and Von Wachter (2019), Piketty (2018) and Piketty and Saez (2003).

References

Acemoglu, D. (2023). Distorted innovation: Does the market get the direction of technology right? In *AEA Papers and Proceedings*, Volume 113, pp. 1–28.

Acemoglu, D. and U. Akcigit (2012). Intellectual property rights policy, competition and innovation. *Journal of the European Economic Association 10*(1), 1–42.

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hémous (2019). Innovation and top income inequality. *Review of Economic Studies 86*(1), 1–45.

Aghion, P., C. Antonin, and S. Bunel (2021). *The power of creative destruction*. Harvard University Press.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2019). Missing growth from creative destruction. *American Economic Review 109*(8), 2795–2822.

[17] See Aghion, Antonin, and Bunel (2021, chapter 12) for a description of the French system of R&D subsidies.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2023). A theory of falling growth and rising rents. *Review of Economic Studies forthcoming*.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and innovation: An inverted-u relationship. *Quarterly Journal of Economics 120*(2), 701–728.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001). Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies 68*(3), 467–492.

Aghion, P., C. Harris, and J. Vickers (1997). Competition and growth with step-by-step innovation: An example. *European Economic Review 41*(3-5), 771–782.

Akcigit, U. and S. T. Ates (2019). What happened to u.s. business dynamism? NBER working paper 25756.

Akcigit, U., S. Baslandze, and F. Lotti (2018). Connecting to power: political connections, innovation, and firm dynamics. NBER working paper 25136.

Akcigit, U., M. A. Celik, and J. Greenwood (2016). Buy, keep, or sell: Economic growth and the market for ideas. *Econometrica 84*(3), 943–984.

Akcigit, U., D. Hanley, and S. Stantcheva (2022). Optimal taxation and r&d policies. *Econometrica 90*(2), 645–684.

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

Amiti, M., O. Itskhoki, and J. Konings (2019). International shocks, variable markups, and domestic prices. *Review of Economic Studies 86*(6), 2356–2402.

Arundel, A. and I. Kabla (1998). What percentage of innovations are patented? empirical estimates for european firms. *Research policy 27*(2), 127–141.

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

Ayerst, S. (2023). Innovator heterogeneity, r&d misallocation and the productivity growth slowdown.

Babalievsky, F. (2022). Misallocation in the market for inventors.

Banholzer, N., V. Behrens, S. Feuerriegel, S. Heinrich, C. Rammer, U. Schmoch, F. Seliger, and M. Wörter (2019). Knowledge spillovers from product and process inventions in patents and their impact on firm performance. end report. Technical report, ZEW-Gutachten und Forschungsberichte.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *Quarterly Journal of Economics 135*(1), 105–163.

Bergeaud, A., A. Guillouzouic, E. Henry, and C. Malgouyres (2022). From public labs to private firms: magnitude and channels of r&d spillovers. Technical Report 1882, Centre for Economic Performance - LSE.

Bils, M. and P. J. Klenow (2004). Some evidence on the importance of sticky prices. *Journal of political economy 112*(5), 947–985.

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

Boar, C. and V. Midrigan (2022). Markups and inequality. NBER working paper 25952.

Cavenaile, L., M. A. Celik, and X. Tian (2021). Are markups too high? competition, strategic innovation, and industry dynamics. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3459775.

Cohen, W. M. and R. C. Levin (1989). Empirical studies of innovation and market structure. *Handbook of industrial organization 2*, 1059–1107.

Cohen, W. M., R. Nelson, and J. P. Walsh (2000). Protecting their intellectual assets: Appropriability conditions and why us manufacturing firms patent (or not).

Dasgupta, P. and J. Stiglitz (1980). Industrial structure and the nature of innovative activity. *Economic Journal 90*(358), 266–293.

Davison, C. (2022). Knowledge spillovers of products and processes. working paper.

De Loecker, J. and J. Eeckhout (2018). Global market power. NBER working paper 24768.

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

De Ridder, M. (2019). Market power and innovation in the intangible economy. working paper 1931, Faculty of Economics, University of Cambridge.

De Ridder, M., B. Grassi, and G. Morzenti (2022). The hitchhiker's guide to markup estimation. working paper.

Edmond, C., V. Midrigan, and D. Y. Xu (2015). Competition, markups, and the gains from international trade. *American Economic Review 105*(10), 3183–3221.

Edmond, C., V. Midrigan, and D. Y. Xu (2021). How costly are markups? NBER working paper 24800.

Farhi, E. and F. Gourio (2018). Accounting for macro-finance trends: Market power, intangibles, and risk premia. *Brookings Papers on Economic Activity* (147), 147.

Fitzgerald, D. and S. Haller (2014). Pricing-to-market: evidence from plant-level prices. *Review of Economic Studies 81*(2), 761–786.

Gopinath, G., O. Itskhoki, and R. Rigobon (2010). Currency choice and exchange rate pass-through. *American Economic Review 100*(1), 304–36.

Gopinath, G. and R. Rigobon (2008). Sticky borders. *Quarterly Journal of Economics 123*(2), 531–575.

Hall, B. H. and V. Sena (2017). Appropriability mechanisms, innovation, and productivity: evidence from the uk. *Economics of Innovation and New Technology 26*(1-2), 42–62.

Hallak, J. C. and P. K. Schott (2011). Estimating cross-country differences in product quality. *The Quarterly journal of economics 126*(1), 417–474.

Haltiwanger, J., R. Kulick, and C. Syverson (2018). Misallocation measures: The distortion that ate the residual. NBER working paper 24199.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016). Quantifying the Sources of Firm Heterogeneity. *Quarterly Journal of Economics 131*(3), 1291–1364.

Klette, T. J. and S. S. Kortum (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy 112*(5), 986–1018.

Lehr, N. H. (2022). Did r&d misallocation contribute to slower growth.

Levin, R. C., A. K. Klevorick, R. R. Nelson, S. G. Winter, R. Gilbert, and Z. Griliches (1987). Appropriating the returns from industrial research and development. *Brookings papers on economic activity 1987*(3), 783–831.

Liu, E. and S. Ma (2021). Innovation networks and r&d allocation. Working Paper w29607, National Bureau of Economic Research.

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

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

Ma, Y. (2021). Specialization in a knowledge economy.

Moser, P. (2005). How do patent laws influence innovation? evidence from nineteenth-century world's fairs. *American Economic Review 95*(4), 1214–1236.

Moser, P. (2013). Patents and innovation: evidence from economic history. *Journal of Economic Perspectives 27*(1), 23–44.

Moulton, B. (2018). The measurement of output, prices, and productivity: What's changed since the boskin commission? *The Hutchins Center on Fiscal and Monetary Policy at the Brookings Institution*.

Ornaghi, C. (2006). Spillovers in product and process innovation: Evidence from manufacturing firms. *International Journal of Industrial Organization 24*(2), 349–380.

Piketty, T. (2018). *Capital in the twenty-first century*. Harvard University Press.

Piketty, T. and E. Saez (2003). Income inequality in the united states, 1913–1998. *Quarterly journal of economics 118*(1), 1–41.

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

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

Voronina, M. (2022). Endogenous growth and optimal market power.

A Robustness checks of the empirical patterns

Table 6: Robustness for bivariate regression of hours on firm TFPQ and price

	(1)	(2)	(3)	(4)	(5)	(6)
TFPQ	0.418	0.290	0.529	0.674	0.372	0.402
	(0.056)	(0.050)	(0.059)	(0.077)	(0.047)	(0.056)
Price	0.472	0.330	0.580	0.725	0.396	0.451
	(0.061)	(0.057)	(0.063)	(0.080)	(0.053)	(0.062)

Notes: This Table presents various robustness checks around the estimation of the relationship between the logarithm of firm hours and price and TFPQ. Column 1 reports the baseline from Table 1. Column 2 adds a control for age interacted with a sector FE. Column 3 measures TFPQ with hours in the denominator and lags it by one year. Column 4 measures TFPQ with hours and instruments it with the baseline measure of TFPQ. Column 5 uses 5 digits sector FE instead of 2 digit sector FE. Column 6 measures the dependent variable using only production workers.