R&D Scale, Markups, and Productivity Growth*

Domenico Ferraro Soroush Ghazi Pietro F. Peretto

February 26, 2024

Abstract

Total factor productivity (TFP) growth rates display a hump-shaped pattern from the mid-1970s onward, peaking in the early 2000s. R&D as a share of firms' sales rises throughout the same period. Motivated by these facts, we develop an endogenous growth model in which large dominant firms face a competitive fringe and charge variable markups. We evaluate the model's ability to match the observations and use it to simulate counterfactual scenarios under alternative configurations of innovation technologies. The model reconciles observed trends in productivity growth and R&D by steadily increasing overhead labor costs of R&D labs. The model also generates increasing profits-to-sales ratios consistent with firm-level data from Compustat.

JEL Classification: D24; E23; O31; O47.

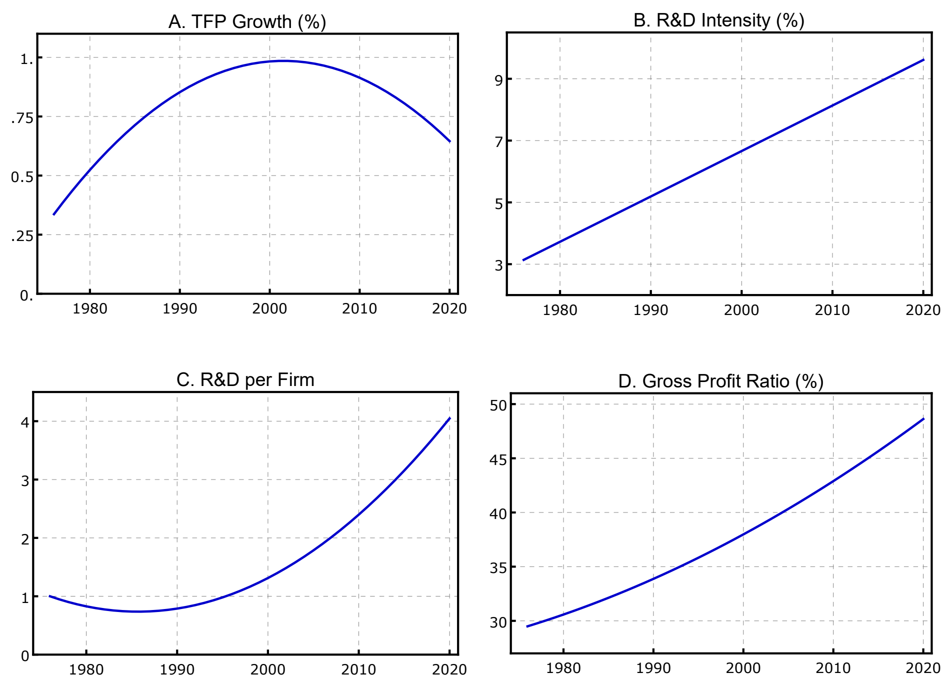
Keywords: Productivity growth; Technological change; Research and Development; Variable markups; Overhead costs.

*Ferraro: Department of Economics, W. P. Carey School of Business, Arizona State University, PO Box 879801, Tempe, AZ 85287, United States (e-mail: domenico.ferraro@asu.edu); Ghazi: Department of Economics, Finance and Legal Studies, Culverhouse College of Business, The University of Alabama, Box 870224, Tuscaloosa, AL 35487, United States (e-mail: sgkalahroudi@ua.edu); Peretto: Department of Economics, Duke University, Box 90097, Durham, NC 27708, United States (e-mail: peretto@econ.duke.edu).
First version: February 15, 2024.

1. Introduction

This paper is motivated by two observations on trends in the postwar United States (see Figure 1). The first observation is that total factor productivity (TFP) growth rates display a hump-shaped pattern from the mid-1970s onward, with a peak in the early 2000s. The second observation is that research and development (R&D) intensity measured as a share of firms' sales and R&D per firm have steadily increased throughout the same period. Coincidentally, gross profit ratios measured as firms' profits to sales have risen.

Figure 1: U.S. Trends in R&D and TFP Growth, 1976-2020



Notes: The figure shows quadratic trends for utilization-adjusted TFP growth rates from Fernald (2014) (panel A), R&D as a share of firms' sales (panel B), R&D per firm, normalized to one in 1976 (panel C), gross profit ratio measured as one minus the ratio of cost of goods sold (COGS) to sales (panel D). Panels B, C, and D consider a sample of Compustat's top-20 highest R&D industries. See Appendix A for more details on data sources, variables' definitions, and construction.

These facts have spurred debate among academics, policymakers, and practitioners. Conceptually, they are puzzling as they require reconciling decreasing innovative output proxied by declining TFP growth with an ever-increasing innovative input: increasingly more resources must be allocated to R&D to preserve a given productivity growth rate, or, put differently, measured research productivity is falling (Bloom, Jones and Van Reenen, 2020; Cowen, 2011; Gordon, 2017).

In this paper, we propose a novel theory based on the minimum efficient scale of R&D labs, quantify it in a new endogenous growth model with variable markups, and show that the model is successful in reproducing the secular trends in TFP growth and R&D intensity and their comovement. Using data and the quantitative theory, we argue that an increasing scale of R&D activities resulting from rising overhead costs of R&D labs can quantitatively account for the patterns in the data. Developing a quantitative theory that successfully reconciles the observed trends is the paper's main contribution, which we relate to the literature in Section 2.

We also use the model to evaluate two competing theories of the observed trends. First, we ask whether a steady decline in research productivity can account for the facts. Our model answer is negative. Declining research productivity leads to counterfactual implications, such as decreasing R&D intensity and declining gross profit ratios of R&D-intensive firms. Both such implications are at odds with microdata from Compustat. Hence, a steady decline in research productivity cannot be the main driving force of the observed trends.

Second, increasing fixed operating costs, like overhead labor, is another candidate driving force behind the observed trends. We quantitatively evaluate and rule out this hypothesis on the grounds of generating a counterfactual falling R&D intensity, again, at odds with microdata from Compustat.

Providing more details of the environment we set up in Section 3 is helpful to grasp insight from our results. The model features endogenous growth fueled by investments by dominant firms facing a competitive fringe and charging variable markups. Dominant firms and their competitive fringes produce intermediate goods imperfectly substitutable in final good production. Dominant firms make innovative investments, allocating labor to reduce unit costs of production and, thereby, contribute to advancing frontier technology. Such firms charge a markup over the marginal cost of production; the markup is increasing in their market share. By contrast, atomistic firms in the competitive fringe do not invest; instead, they tap into publicly available technology competing with dominant firms; they make zero profits, pricing at marginal costs.

The equilibrium growth rate of technology results from two offsetting forces: the *cost-spreading* effect versus the *business-stealing* effect. The former positively affects the growth rate as the market share of large firms increases, while the latter negatively affects it.

First, the cost-spreading effect relates to the scale or market-size effect at the core of R&D-based growth models. The return to R&D is increasing in the scale of operation and, hence, in dominant firms' market share. The intuition for the positive effect of cost-spreading is that a dollar spent on R&D lowers the cost for every unit produced, implying that the gains of R&D are proportional to the scale of operation.

Second, the business-stealing effect emerges from dominant firms internalizing their impact on the price of the product line. The intuition is that the demand for a dominant firm's product depends on the price it charges relative to the price level of the product line. However, as R&D lowers the price of the dominant firm, it also reduces the price of the product line, and the latter lowers the return to R&D, thereby disincentivizing R&D. Such an effect strengthens as the market share of the dominant firm rises since the aggregate price index moves in lockstep with the dominant firm's price. Effectively, it becomes harder to steal the market share of other firms by doing R&D as one's market share increases.

These opposing forces can induce a hump-shaped technology growth rate profile in dominant firms' market share, thus reconciling the trend in the data. Early on, when the market share of dominant firms is small, the positive effect of cost-spreading dominates, and the technology growth rate increases in the market share; however, after the market share reaches a critical threshold, the business-stealing effect turns to dominate, and the technology growth rate decreases. Quantifying the relative strength of cost-spreading and business-stealing is an interesting empirical question our model is well suited to address.

We analytically characterize the comparative statics of the technology growth rate for the research productivity and cost parameters, assuming that the economy is in the steady state in Section 4. The research productivity parameter governs the marginal productivity of labor in R&D, shifting the firm's knowledge production possibility frontier akin to the technology parameter in the neoclassical growth model. The overhead cost parameters influence the efficient minimum output and knowledge production scale. Three main results stand out.

First, if the cost of R&D increases due to a decline in the productivity of R&D labor, the dominant firms' market share and profitability drop, together with their R&D intensity, all of which counterfactual to their respective trends in U.S. microdata from Compustat. A lower firm's productivity in knowledge creation reduces the return to R&D, which lowers their R&D intensity and technology growth rate. At the aggregate level, resources shift from R&D to firm creation, lowering dominant firms' market shares and profit ratio.

In contrast, if the cost of R&D increases due to an increase in the overhead costs of R&D, the market share of large firms, their profit ratio, and R&D intensity increase, while the aggregate growth rate follows a hump-shaped pattern, consistent with the data. The intuition is that an increase in overhead costs does not directly affect the productivity of R&D labor; however, it lowers dominant firms' distributed dividends and, thereby, the return to entry. Reduced entry, in turn, implies an increase in the market share of dominant firms, which induces hump-shaped aggregate technology growth. Further, a larger market share implies a higher profit ratio. As the firm's minimum scale of R&D labs increases, more resources are allocated away from output production to R&D, inducing an increasing R&D intensity path. Again, this prediction nicely squares with the observed time series of R&D intensity in the data.

Finally, we show that rising overhead costs of production, as opposed to overhead costs of R&D, cannot be the sole driver of the trends in the data. An increase in the overhead cost of production raises the scale of output production, which shifts the firm's resource allocation away from R&D, resulting in a counterfactual R&D intensity pattern.

In Section 5, we quantify these model's predictions and contrast them with U.S. data. An appealing feature of the model is that the general equilibrium is highly tractable in and out of the steady state. A second-order differential equation fully describes transition dynamics in dominant firms' market share. Once one solves for the equilibrium time path of such market share, allocations and prices are easily calculated. This property allows for a transparent quantification of the mechanisms behind the observed trends in the data.

To operationalize the idea that rising overhead costs of R&D drive the observed trends, we construct a measure of "patent complexity" based on the number of technological classes spanned by granted U.S. patents, using data from the USPTO (U.S. Patent and Trademark Office). Consistent with the view in Jones (2009) that advancing technology requires increasingly specialized knowledge, we find that the average number of CPC (Cooperative Patent Class) classifications cited by patents issued to U.S. corporations has steadily risen over time. Lacking direct measurement of overhead costs, we use time series variation in such patent complexity measurement to discipline the time path of the overhead cost parameters in the model. The rationale is that increasing specialization goes hand in hand with the observed rising research team sizes (Jones, 2021).

Hence, we do not back out the overhead costs by requiring the model to match the observed trends we aim to explain; instead, we use arguably extraneous information to inform the time variation in the overhead cost parameters of R&D we feed to the model. The implied rising overhead costs of R&D effectively increase the scale of R&D labs, thus requiring increasingly more labor so that firms must allocate increasingly more resources to R&D to preserve a given technology growth rate. The calibrated model successfully reproduces the hump-shaped pattern of TFP growth rates from the late 1970s onward with the steadily increasing R&D intensity and gross profit ratios as in the data. In the model and consistent with the data, measured research productivity falls over time.

We structure the paper as follows. In Section 2, we discuss closely related papers and highlight our paper's contributions. We present the model in Section 3 and discuss our main analytical results in Section 4. In Section 5, we take the model to the data and quantify its implications for the U.S. economy. Section 6 concludes.

2. Related Literature

The paper contributes a novel mechanism to reconcile the puzzling disconnect between rising innovative inputs such as R&D investments and falling innovative output proxied by TFP growth. In addition to the work already cited, this section discusses the closely related papers by Aghion et al. (2023), De Ridder (2024), and Olmstead-Rumsey (2022). Despite similarities, we will highlight the crucial differences between our work and theirs.

Aghion et al. (2023) develop a model in which reduced overhead costs of operating multiple product lines drive declining productivity growth and rising concentration. In their model, as such overhead costs fall, more efficient firms expand their span of control into new product lines, thereby increasing concentration. In the early phase, increased concentration spurs a productivity growth acceleration followed by a slowdown when incentives to innovate taper off. Reduced overhead costs of operating multiple product lines raise incentives for innovative investments. However, such an effect is offset and, under some parameter values, overturned by lower expected markups due to increased competition.

De Ridder (2024) develops a model that reconciles the observed trends of productivity growth and R&D intensity in the U.S. since the mid-1980s by the sustained increased use of intangible inputs in production, such as software. Rising intangible inputs shift a firm's cost structure towards fixed costs, breaking the positive link between R&D and productivity growth. As firms with intangible intensive production enter the market, the economy experiences a productivity growth acceleration associated with higher markups. After this early phase, a slowdown ensues as firms with high intangibles deter entry. At the same time, incumbents with low intangibles face reduced incentives to innovate due to harsher competition.

Finally, Olmstead-Rumsey (2022) provides novel evidence based on patent data and a quantitative model to argue that the laggard firms' reduced ability to compete with leader firms through innovation can account for a significant share of the productivity slowdown with rising concentration.

What distinguishes our work from the three papers above is twofold.

First, we develop a new quantitative yet highly tractable, fully endogenous growth model with variable markups, contributing a novel mechanism. The model features markups that are a function of market share, combining two dimensions of innovation: firm entry that expands product variety and cost-reducing investments by incumbent dominant firms that expand the frontier technology. Horizontal and vertical innovation play different but equally important roles: firms' investments reducing unit production costs allow for endogenous growth in the long run; expanding variety makes the model's steady-state growth rate scale-free (Peretto, 1998; Peretto and Connolly, 2007; Peretto and Smulders, 2002). Importantly, population size has no bearing on the productivity growth rate in the long run, which enables us to conceptually isolate the effect of the rising scale of R&D labs from the impact of varying population size on productivity growth, which is the focus of Hopenhayn, Neira and Singhania (2022) and Karahan, Pugsley and Şahin (2019).

Second, we use the model to discriminate among alternative competing explanations of the observed trends. We quantify that the rising scale of R&D activities due to rising—not falling—overhead costs of R&D labs goes a long way in reconciling several observed trends in productivity growth, R&D intensity, and concentration. To our knowledge, this is the first paper making this point.

In addition, and relatedly, we show that such a rise in the scale of R&D labs correlates with the increasing complexity of patents, which we measure as the technological span of granted patents. In our quantitative exercise, we do not hard-wire overhead costs to rising concentration; instead, we discipline them by somewhat extraneous information on patent complexity, leaving the model free to deliver the observed trends in the data. The model generates the observed hump-shaped pattern of TFP growth even though the implied overhead costs of R&D we back out from the data rose steadily throughout the period.

3. Model

The model environment encompasses a representative household, a final good producer, intermediate good producers, and entrants that launch new products. The model has no physical capital, and the household inelastically supplies labor needed in all production processes, including product creation.

The production side of the economy consists of a competitive final good producer that buys differentiated intermediate goods as input. Each intermediate good is the output of a product line with a dominant firm and competitive fringe market structure. The fringe produces a good that is an imperfect substitute for the dominant firm's product. In each product line, the dominant firm has a substantial market share and, hence, market power, but the fringe are atomistic small firms that individually have no market power and act as price takers.

There are two means of innovation. First, dominant firms invest in their technology to reduce marginal costs. Second, entrepreneurs can invest and initiate a new product line as the dominant firm.

3.1. Final Good Producers

A representative final good producer buys intermediate goods produced in all product lines by the dominant firms and their fringes, combines them into the final good, and sells it competitively. We denote the final good with $\mathcal{X}$ and index each product line with i. The dominant firm in the product line i produces x_i, and the small firms in its fringe collectively produce x_{0i}. The final good producer combines x_i and x_{0i} into X_i using a CES technology and then combines the X_i's using a generalized Cobb-Douglas technology to produce $\mathcal{X}$,

$$\ln \mathcal{X} = \frac{1}{N}\int_0^N \ln X_i \, di, \quad \text{with} \quad X_i = \left(x_i^{\frac{\epsilon-1}{\epsilon}} + \omega \, x_{0i}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}, \quad \epsilon > 1, \tag{1}$$

where N is the number of product lines.

The generalized Cobb-Douglas production function in (1) implies that the final good producer spends equally on each product line i, which induces the following demand schedule for X_i,

$$X_i = \left(\frac{\mathcal{P}\mathcal{X}}{N}\right)\frac{1}{P_i}. \tag{2}$$

In (2) , $\mathcal{P}$ denotes the price of the final good, and P_i is the (shadow) price of X_i. Denoting the expenditure on the final good with $\mathcal{Y}$, and that on the product line i with Y_i, we have

$$Y_i = \frac{\mathcal{Y}}{N}, \quad \text{where} \quad Y_i \equiv P_i X_i, \quad \text{and} \quad \mathcal{Y} \equiv \mathcal{P}\mathcal{X}. \tag{3}$$

The CES technology implies that the demand for the dominant firm i's product x_i and its competitive fringe's product $x_{0,i}$ are

$$x_i = \left(\frac{p_i}{P_i}\right)^{-\epsilon} X_i, \quad \text{and} \quad x_{0i} = \left(\frac{p_{0i}}{\omega P_i}\right)^{-\epsilon} X_i, \tag{4}$$

where p_i and p_{0i} are the prices of x_i and x_{0i} respectively. In terms of expenditure,

$$x_i = Y_i \frac{p_i^{-\epsilon}}{P_i^{1-\epsilon}}, \quad \text{and} \quad x_{0i} = \omega^{\epsilon} Y_i \frac{p_{0i}^{-\epsilon}}{P_i^{1-\epsilon}}. \tag{5}$$

Perfect competition in the final good sector implies

$$P_i^{1-\epsilon} = p_i^{1-\epsilon} + \omega^{\epsilon} p_{0i}^{1-\epsilon}. \tag{6}$$

For future reference, we denote the market share of the dominant firm i in product line i with s_i so that

$$s_i \equiv \frac{p_i x_i}{P_i X_i} = \frac{p_i \left(\frac{p_i}{P_i}\right)^{-\epsilon} X_i}{P_i X_i} = \frac{p_i^{1-\epsilon}}{P_i^{1-\epsilon}}. \tag{7}$$

3.2. Dominant Firms

The productivity of the dominant firm i depends on its own technology level z_i and the level of publicly available technology Z.[1] Public technology is an increasing function of private technologies and captures the positive effect of knowledge spillovers, arguably the most crucial externality regarding R&D (see, e.g., Jones and Williams, 2000; Hall,

[1]We use the terms knowledge and technology interchangeably.

Mairesse and Mohnen, 2010). Specifically, a dominant firm's technology is

$$x_i = z_i^{\theta} Z^{1-\theta} (\ell_{x,i} - \psi), \tag{8}$$

where $\ell_{x,i}$ is labor that the dominant firm i allocates to goods production, and ψ is the fixed operating cost of production. In (8), private and public technologies complement each other in production (Cohen and Levinthal, 1989; Ferraro, Ghazi and Peretto, 2023).

We define the level of public technology, Z, as

$$Z \equiv \frac{\int_0^N z_j \, dj}{L}, \tag{9}$$

which is increasing in the stock of private knowledge of individual dominant firms, z_j, and the mass of dominant firms, N. The latter captures the fact that dominant firms produce differentiated goods, possessing different types of knowledge to contribute to the aggregate level of knowledge. The L in the denominator in (9) is population size, which equals to the economy's total labor supply. It ensures that there is no scale effect, i.e., the steady-state productivity growth rate is scale invariant.[2] The intuition is that along the balanced growth path, the number of product lines N is proportional to L, so that from (9), public knowledge in symmetric equilibrium is independent of L.

Each dominant firm i can invest in its technology and increase z_i to reduce the marginal cost of production. To increase z_i, a dominant firm has to hire labor $\ell_{z,i}$, which is the sole input into the knowledge production function,

$$\dot{z}_i = \alpha Z (\ell_{z,i} - \phi), \tag{10}$$

where α is the research productivity parameter governing the marginal cost of creating new knowledge or ideas.[3] We also assume that running an R&D lab is subject to a fixed operating cost, ϕ, akin to the formulation of the fixed operating cost in the technology of output production, ψ, in (8). In this sense, output and knowledge production are treated symmetrically. More specifically, ϕ regulates the minimum number of scientists needed to run the lab, a feature of the knowledge creation technology that is taken parametrically by the individual firm. For example, creating a new technology that draws from multiple

[2] Any public knowledge aggregator of the form $Z = \frac{\int_0^N z_j \, dj}{N^{\nu} L^{1-\nu}}$, with $\nu \in [0, 1]$, accomplishes neutralization of the scale effect of population size on the steady-state growth rate of technology. Here, we continue with the simpler aggregator (9) to streamline the model.

[3] More generally, one can specify $\dot{z}_i = \alpha z_i^{\eta} Z^{1-\eta} (\ell_{z,i} - \phi_t)$ without any substantive change to the results.

fields needs expertise from all those fields, increasing the minimum scale of R&D labs.

Each dominant firm i takes the final good demand as given and chooses its price and R&D labor to maximize the market value, v_i, i.e., the present discounted value of its future dividends,

$$v_i \equiv \int_t^{\infty} e^{-\int_t^{\tau}(r(s)+\delta)\,ds} d_i(\tau)\,d\tau,$$

where r denotes the interest rate, δ is the exogenous and constant probability of product obsolescence, and d_i is the dividend flow. At each point in time, d_i is revenues or sales net of all costs. We take labor as the numéraire and normalize the wage to one ($W \equiv 1$). The total cost of the dominant firm i is the sum of wages paid to the labor in production and R&D, $\ell_{x,i} + \ell_{z,i}$. Thus, the dominant firm's dividend flow is $d_i = p_i x_i - (\ell_{x,i} + \ell_{z,i})$.

The demand schedule $x_i = Y_i \frac{p_i^{-\epsilon}}{P_i^{1-\epsilon}}$ implies the price elasticity

$$e_i = \epsilon - (\epsilon - 1)s_i, \tag{11}$$

which, in turn, implies the endogenous price markup

$$p_i = \frac{e_i}{e_i - 1} \cdot \frac{1}{z_i^{\theta} Z^{1-\theta}}, \tag{12}$$

where $1/(z_i^{\theta} Z^{1-\theta})$ is the marginal cost of producing a unit of x_i. Given e_i, the higher the firm's technology level, z_i, the lower the marginal cost of production. Similarly, the higher the public knowledge level, Z, the lower the firm's marginal cost. Such cost-reducing effects of firm and economy-wide knowledge or technology levels are the key drivers of firms' R&D labor allocations and the economy's long-run productivity growth.

Note also that the price elasticity e_i is decreasing in the market share s_i, and the lower e_i, the higher the price markup over the marginal cost. Thus, markups and the dominant firm's market share are positively related.

Dominant firm's R&D problem is dynamic and implies the following expression for the rate of return to R&D, whose derivations we detail in Appendix B,

$$r = -\delta - \frac{\dot{Z}}{Z} + \alpha\theta \frac{Z}{z_i} \cdot Y_i s_i \cdot \frac{e_i - 1}{e_i}, \text{ if } \ell_{z,i} > 0. \tag{13}$$

If the return to R&D is sufficiently low, dominant firms choose not to invest, $\ell_{z,i} = 0$, and the right-hand side of (13) is less than r.

Inspection of (13) reveals that the term $-\frac{\dot{Z}}{Z} - \delta$ reduces the rate of return to R&D. A higher growth rate of public knowledge reduces the shadow value of private technology, thereby lowering the rate of return of allocating an additional unit of labor to R&D away from output production.[4] Similarly, a higher rate of product obsolescence reduces the expected lifetime of a product, thus disincentivizing R&D.

The last term, $\alpha\theta\frac{Z}{z_i} \cdot Y_i s_i \cdot \frac{e_i - 1}{e_i}$ consists of three parts. First, $Y_i s_i$ captures the *cost-spreading* effect that is increasing in the market size served by the dominant firm, that is, the dominant firm i's market share s_i of total expenditure on the product line, Y_i. The larger the scale of operations, the higher the return to R&D.

Second, $\frac{e_i - 1}{e_i}$ captures the *business-stealing* effect that is decreasing in the market share due to the dominant firm internalizing its impact on the price of the product line P_i.[5] The intuition is that a dominant firm that invests in R&D reduces its price to capture the market share from its fringe—moreover, increasing R&D results in a more significant capture of market share if the firm has a smaller market share.

Third, $\frac{Z}{z_i}$ captures the role of *technology spillovers*, which negatively depends on the dominant firms' market share in symmetric equilibrium, magnifying the impact of the cost-spreading and business-stealing effects.

3.3. Competitive Fringe

In each product line, atomistic competitive firms produce a homogeneous good that is an imperfect substitute for the dominant firm's product. As firms in the competitive fringe make zero profits, they cannot invest and advance their private technology. However, the competitive fringe has access to the publicly available technology Z, which, in turn, determines their marginal cost.

Specifically, a competitive fringe firm's technology is

$$x_{0i} = Z\ell_{x0,i}. \tag{14}$$

Perfect competition then implies that the fringe prices at the marginal cost,

$$p_{0i} = \frac{1}{Z}. \tag{15}$$

[4]The shadow value of the firm's private technology or the costate variable of the Hamiltonian of the firm's R&D dynamic problem is $\lambda = 1/(\alpha Z)$. The growth rate of such shadow value is then $\dot{\lambda}/\lambda = -\dot{Z}/Z$.

[5]When a large firm does not internalize its effect on the price, the expression reduces to $\frac{\epsilon - 1}{\epsilon}$ as if the firm has zero market share.

For future reference, note that the "number" of firms in the fringe is inconsequential for allocations and prices. Setting up a firm on the competitive fringe is costless. Thus, the product line is immediately populated by a competitive fringe when a new dominant firm launches a new product.

3.4. Firm Entry and Product Creation

Launching a new product requires incurring a sunk entry cost. Since a competitive fringe immediately emerges, firm entry leads to a new product line. We assume that the cost of launching a new product is proportional to the gross profit flow of a dominant firm.[6] Free-entry then implies that the value of a dominant firm i creating a new product is

$$v_i = \gamma \frac{sY}{e}. \tag{16}$$

In (16), the right-hand side variables do not have a subscript to emphasize that a new dominant firm enters at the average values.[7] Since labor is the only factor of production, the entry cost can be thought as the labor needed to create a dominant firm.

The rate of return to entry is subject to the familiar no-arbitrage condition,

$$r = \frac{d_i + \dot{v}_i}{v_i} - \delta, \tag{17}$$

for positive entry.

3.5. Households

There is an infinitely-lived representative household facing the standard consumption-saving problem. Household's preferences over deterministic sequences of consumption $\{C_t\}_{t=0}^{\infty}$ are described by the utility function $U(C_t) = \ln(C_t)$, where C_t is consumption at date $t \geq 0$. The household's problem takes the familiar form:

$$\max_{\{C_\tau\}_{\tau=0}^{\infty}} \int_t^{\infty} e^{-\rho(\tau - t)} \ln(C_\tau)\, d\tau \quad \text{s.t.} \quad \dot{Q} = rQ + L - \mathcal{P}C, \quad \forall \tau \geq t, \tag{18}$$

[6]Appendix Section A.2.1 details how to calculate the model-implied long-run value of the firm by using data on R&D-to-Sales and COGS-to-Sales ratios from Compustat.

[7]The gross profit flow is $px - (l_x - \psi)$, where $\ell_x - \psi$ is the cost of goods sold (COGS) and ψ is the overhead cost. Further, $l_x - \psi = x \cdot \text{marginal cost} = px(e-1)/e$ so that $px - (l_x - \psi) = px/e = sY/e$.

where ρ is the time discount rate and Q denotes the household's financial wealth, the state variable of the household's problem. Standard first-order conditions give the familiar Euler equation,

$$r = \frac{\dot{\mathcal{Y}}}{\mathcal{Y}} + \rho, \tag{19}$$

where $\mathcal{Y} \equiv \mathcal{P}C$ is total consumption expenditures.

3.6. General Equilibrium

The equilibrium is a time path of allocations $\{\mathcal{X}, C, X_i, x_i, x_{0i}, \ell_{x,i}, \ell_{z,i}, \ell_{x0,i}, L_N\}$, and prices $\{\mathcal{P}, P_i, p_i, p_{0i}, v_i, r\}$, such that given the state variables $\{N, z_i\}$, the allocations and prices solve the problems of the final good producer, dominant firms and their competitive fringe, entrepreneurs, and households. Moreover, product and labor markets clear so that $C = \mathcal{X}$ and $L = \int_0^N (\ell_{x,i} + \ell_{z,i} + \ell_{x_{0i}})\, di + L_N$, where L_N denotes total labor in new product creation.

The following proposition allows us to work with the model's symmetric equilibrium.

Proposition 1. *The equilibrium of the model is symmetric and the market share of large firms can take any initial value larger than* $\underline{s} \equiv \frac{2\epsilon - 1 - \sqrt{1+4\epsilon/\theta}}{2(\epsilon-1)}$.

Proof. See Appendix B. □

In Proposition 1, if the initial technology level z_i of a dominant firm is sufficiently high (relative to the public technology), such that its initial market share is higher than $\underline{s}$, then its market share converges to the average dominant firms' market share. Otherwise, if a dominant firm starts with a sufficiently low private technology, it never catches up with the average dominant firm in the model and eventually drops out. As the model supports a symmetric equilibrium, we drop the subscript i from now on.

Solving for the general equilibrium requires a few steps. First, starting with the budget constraint (18), we use product market clearing, $\mathcal{Y} = \mathcal{P}C$, and divide through by Q to obtain $\dot{Q}/Q = r + (L - \mathcal{Y})/Q$. Then, using asset market clearing, $Q = Nv$, and the free-entry condition (16) so that $r + (L - \mathcal{Y})/Q = \dot{\mathcal{Y}}/\mathcal{Y} + \dot{s}/s - \dot{e}/e$, we substitute $r = \dot{\mathcal{Y}}/\mathcal{Y} + \rho$ from the household's Euler equation (19) and solve for

$$\mathcal{Y} = \frac{L}{1 - \gamma \frac{s}{e}(\rho + \frac{\dot{e}}{e} - \frac{\dot{s}}{s})}, \tag{20}$$

which gives final good expenditure as a function of solely the market share of dominant firms, s. (The price elasticity e only depends on s.) Finally, Euler equation (19) gives the interest rate r as a function of s given $\mathcal{Y}$.

Second, we establish that the measure of products, $n \equiv N/L$, is decreasing in s. In symmetric equilibrium, public knowledge (9) equals $Z = nz$ so that the market share of dominant firms in (7) can be rewritten as

$$s = \frac{p^{1-\epsilon}}{p^{1-\epsilon} + \omega^{\epsilon} p_0^{1-\epsilon}} = \frac{1}{1 + \omega^{\epsilon} \left(\frac{p}{p_0}\right)^{\epsilon - 1}} = \frac{1}{1 + \omega^{\epsilon} \left(\frac{e}{e-1} n^{\theta}\right)^{\epsilon-1}},$$

which solving for n gives

$$n = \omega^{\frac{-\epsilon}{(\epsilon-1)\theta}} \left(\frac{e-1}{e}\right)^{\frac{1}{\theta}} \left(\frac{1-s}{s}\right)^{\frac{1}{(\epsilon-1)\theta}}. \tag{21}$$

Corollary 2. *The measure of products, n, is strictly decreasing in the market share of dominant firms, s.*

Proof. See Appendix B. $\square$

The link between the measure of products in the economy and the market share of each dominant firm in its own product line stems from technology spillovers. Specifically, a smaller n lowers the ratio of public to private technology, which makes the firms in the fringe less competitive, causing them to lose market share.

Third, we rewrite (13) to obtain the dominant firms' technology growth rate as

$$\mathfrak{z} \equiv \dot{z}/z = -(r+\delta) - \frac{\dot{n}}{n} + \alpha\theta \frac{\mathcal{Y}}{L} \cdot s \cdot \frac{e-1}{e}, \quad \text{if } \mathfrak{z} > 0, \tag{22}$$

where the endogenous variables on the right hand side (r, n, $\mathcal{Y}$, and e) are functions of s.

Fourth, it remains to calculate the equilibrium trajectory of the dominant firm's market share, $s(t)$, for all $t \geq 0$. Remaining allocations and prices are then easily calculated given $s(t)$. Starting with the rate of return to entry (17), and using (19)-(22), we obtain (after some algebra) the second-order nonlinear differential equation that pins down $s(t)$,

$$\frac{\dot{n}}{n}\left(1 - \frac{eL}{\alpha\gamma s \mathcal{Y}}\right) + \frac{\dot{e}}{e} - \frac{\dot{s}}{s} = \frac{1 - \theta(e-1)}{\gamma} + \left(r + \delta\right) \frac{eL}{\alpha\gamma s\mathcal{Y}} - \frac{enL}{\gamma s \mathcal{Y}}(\psi + \phi), \tag{23}$$

where the endogenous variables entering the equation are functions of $s(t)$ only.[8]

[8]See Appendix B for derivations where we make explicit that n, e, and $\mathcal{Y}$ are functions of s.

4. Inspecting the Mechanism

This section discusses how the research productivity and overhead cost parameters affect the steady-state equilibrium and the forces that sterilize the "scale effect" of population size on the economy's steady-state growth rate. The model admits a balanced growth path (BGP) where all trending variables grow at a constant (endogenous) growth rate.

Research Productivity (α) and Cost Parameters (ϕ, ψ) Although the quantitative results in the later section concern the transitional dynamics induced by varying overhead costs of R&D labs, here, we provide analytical results assuming the economy has settled on the BGP. We consider the parameters ϕ, α, and ψ, and focus on the comparative statics of the technology growth rate, R&D intensity (R&D-to-Sales ratio), and gross profit ratio (one minus COGS-to-Sales ratio, where COGS refers to the cost of goods sold).

Proposition 3. *On the balanced growth path, the overhead cost of R&D labs (ϕ), the overhead cost of production (ψ), and the research productivity parameter (α), that regulates the marginal productivity of labor in the R&D labs, all positively affect the market share of dominant firms (s), provided that θ and γ are not too large.*

Proof. See Appendix B. □

Proposition 3 implies that a rise in the R&D costs that originates from a decline in the marginal productivity of labor in R&D (a decline in α) decreases the market share of dominant firms and causes them to grow smaller. In contrast, a rise in the cost of R&D that originates from a rise in the fixed operating costs of labs (a rise in ϕ) causes firms to grow larger, which is consistent with the data.

The intuition behind the result is the following. As either of the fixed costs ϕ or ψ increases, dividend flows drop. This implies that the return on entry drops; hence, entry drops, and the measure of products drops. As a result, the market share of dominant firms increases. On the other hand, a decline in α directly lowers the return on R&D investment, making the dominant firms less competitive and losing market share to their competitive fringe.

Building on Proposition 3, the following proposition shows how the three parameters of interest (ϕ, ψ, and α) affect the technology growth rate, $\mathfrak{z}$, on the BGP.

Proposition 4. *Given the conditions in Proposition 3, on the balanced growth path:*

(1) *As the overhead cost of R&D labs* (ϕ) *or the overhead cost of production* (ψ) *rises, the technology growth rate* ($\mathfrak{z}$) *first increases and then declines. That is,* $\mathfrak{z}$ *is hump-shaped in both types of fixed cost parameters,* ϕ *and* ψ*.*

(2) *As* α *decreases, the technology growth rate* ($\mathfrak{z}$) *declines.*

Proof. See Appendix B. □

The first part of Proposition 4 hinges on the fact that neither type of fixed cost can directly change the marginal product of labor in the R&D labs, and hence, they affect the return to R&D only indirectly through changing the market share in general equilibrium. As either of the fixed costs increases, so does s. However, when s is small, the cost-spreading and spillover effects in the rate of return to R&D dominate the business-stealing effect. Thus, the return to R&D is increasing in s. However, after s passes a threshold level, s_0, the business-stealing effect dominates, and the general equilibrium return to R&D decreases. In general equilibrium, the return to R&D determines the technology growth rate in the long run.

The second part of Proposition 4 is more straightforward. A reduction in α directly affects labor productivity in the R&D, lowering the R&D return for any s. This reduces R&D of dominant firms and, thereby, lowers the long-run technology growth rate.

Next, we discuss the effect of the cost parameters on the other two variables of interest: the gross profit ratio (one minus COGS-to-Sales) and R&D intensity (R&D-to-Sales).

Proposition 5. *Given the conditions in Proposition 3, on the balanced growth:*

(1) *As the fixed cost of R&D labs* (ϕ) *or the fixed cost of production* (ψ) *rises, the gross profit ratio increases.*

(2) *As* α *decreases, the gross profit ratio declines.*

Proof. See Appendix B. □

The intuition behind Proposition 5 is that a higher market share of dominant firms (s) is associated with a less elastic demand (smaller e), which leads to more profitable dominant firms. However, from Proposition 3, we know an increase in the fixed costs raises the market share and hence raises the gross profit ratio. In contrast, a smaller α lowers the market share and lowers the gross profit ratio.

The following proposition shows how the cost parameters affect R&D intensity.

Proposition 6. *Given the conditions in Proposition 3, on the balanced growth:*

(1) *As the overhead cost of R&D labs (ϕ) rises, the R&D-to-Sales ratio increases.*

(2) *As α declines, the R&D-to-Sales ratio decreases.*

(3) *The effect of the overhead cost of production (ψ) on the R&D-to-Sales ratio depends on ϕ. If ϕ is zero, the effect goes from positive to negative as s increases, and if ϕ is sufficiently large, the effect is uniformly negative.*

Proof. See Appendix B. □

Firm-level R&D intensity measures the relative size of R&D labs. The overhead cost of R&D regulates their minimum size without directly affecting labor productivity in R&D labs. Thus, an increase in ϕ increases the relative size of R&D labs. In contrast, a decline in α directly lowers labor productivity in R&D labs and makes them grow relatively smaller.

The third part of Proposition 6 hinges on the fact that a higher ψ affects the relative size of R&D labs only through an increase in market share of dominant firms, s. In particular, the business-stealing effect in the rate of return to R&D depends on the inverse markups (i.e., $\frac{e-1}{e}$, which is decreasing in s), as we discussed earlier based on (13). The negative effect of business stealing makes the relative size of R&D labs shrink when s is large.

A substantive lesson from Proposition 6 is that a decline in the research productivity parameter α cannot explain the observed trends in declining technology growth, $\mathfrak{z}$, since it counterfactually implies a decline in the R&D intensity. Moreover, having an increasing overhead cost in output production is not enough. As the third part of Proposition 6 suggests, an increase in ψ produces a counterfactual pattern in the R&D-to-Sales ratio if there is no fixed cost in R&D labs ($\phi = 0$). This is because although the effect could be initially positive as ψ increases, the effect gradually becomes zero and then negative, at odds with the data displaying a strong and slightly accelerating rise in R&D intensity. For $\phi > 0$, the effect of an increase in ψ continues to generate counterfactual implications.

Scale-Free Long-Run Productivity Growth Before turning to the quantitative analysis, it is helpful to stress that the steady-state technology growth rate, $\mathfrak{z}$, does not depend on the scale of the economy, L.

Proposition 7. *There is no scale effect on the balanced growth path so that the steady-state growth rate of technology $\mathfrak{z}$ is independent of population size, L.*

Proof. See Appendix B. □

In the steady state, the mass of products, N, is proportional to population size, L, so their ratio $n \equiv N/L$ is constant. A permanently larger L leads to a permanently higher N, leaving the growth rate of technology unchanged. As a by-product, the market share of a dominant firm, s, becomes independent of L, too. Notably, while technology growth, in the long run, is scale-free and invariant to population size, it turns out to be hump-shaped in s. This model property will be critical for the model to match the trends in the data to which we turn next.

5. Quantitative Implications for the U.S. Economy

In this section, we quantify the extent to which a steadily increasing minimum scale of R&D labs driven by rising overhead labor can jointly account for the observations on TFP growth, R&D intensity, and gross profit ratios (one minus the COGS-to-Sales ratio) from the late 1970s onward in the United States. We first describe data and measurement of the variables of interest, turn to the model's parameterization, and conclude by quantifying the model's predictions.

5.1. Data and Measurement

This subsection briefly describes the firm-level data from Compustat and USPTO patent data used for the model's quantification. Utilization-adjusted TFP growth rates for the U.S. business sector are from Fernald (2014). Appendix A provides further details and additional facts.

R&D Intensity and Gross Profit Ratio To construct the R&D-to-Sales and COGS-to-Sales ratios, we use the "Fundamentals Annual" file from the CRSP-Compustat merged data, which contains the universe of U.S. publicly traded firms. Again, COGS refers to the cost of goods sold and includes all costs and expenses directly related to producing goods. R&D consists of all costs incurred during the year that relate to the development of new products or services and the improvement of existing products or services.

As is well known, a few industries account for most of U.S. R&D (Cohen and Levin, 1989; Cohen, 2010). Here, instead of calculating R&D intensity aggregating over the R&D-to-Sales ratios of all firms across all industries, we select a subset of industries that account for the lion's share of total R&D and construct an average R&D intensity by aggregating over all firms in such industries with sales shares as weights. Appendix A.2.2 shows that twenty 4-digit NAICS industries cover about 90% of R&D in Compustat. Such a figure suggests that the remaining industries only account for a tiny share of the total. We then calculate average R&D-to-Sales and COGS-to-Sales by dividing total R&D and COGS by total sales for the subset of the twenty high R&D industries.

A Measure of "Patent Complexity" Absent direct measurement of overhead costs of R&D operations, we build a measure of patent complexity and use it to discipline the time path of overhead costs in the model. While imperfect, the idea behind this approach is that advancing frontier technology requires increasingly specialized technical knowledge. And increasing specialization in knowledge creation goes hand in hand with an ever-increasing scale of R&D labs. The model in the previous section captures this idea through a rising minimum labor requirement or overhead R&D labor.

We use PatentsView data from the USPTO (U.S. Patent and Trademark Office) and calculate the average number of CPC (Cooperative Patent Class) classifications in patent citations granted to U.S. corporations. PatentsView data started in 1976, the starting year of all our time series. Fleming and Sorenson (2001) show the number of technological classes to which a patent belongs is a reliable measure of patent complexity.[9]

Using patents' CPC counts as a proxy for their complexity has several advantages that make it uniquely appealing for our purpose. USPTO and EPO (European Patent Office) jointly developed CPC, replacing the U.S. Patent Classification (USPC) in 2015. CPC provides more specific and detailed technology classes than the IPC (International Patent Classification) system. The "Current CPC" file of granted patents in PatentsView data contains the most recent classification. As they put it, "the classification of a granted patent may change, for example, when a new class or subclass is added to the schema." Such retroactive features of patent classification are not readily available for other publicly available measures, such as IPC.

We construct the annual time series of the average current CPC for U.S.-granted patents as follows. We first count the number of unique CPC groups (the finest CPC measure) for each patent; then, we average the CPC count for each year in the data. Using the sample of all U.S. (corporate) patents and at the highest granularity level substantially increases our measure's reliability (Benner and Waldfogel, 2008). We consider the average annual

[9] Aharonson and Schilling (2016) provide crosswalks from patent classifications to technological classes.

CPC count (CPC count for short) as our aggregate patent complexity measure.

5.2. Parameterization

Bringing the model to the data involves two steps. In the first step, we estimate smooth trends by fitting quadratic polynomials to the time series of patent complexity measured by the CPC count, R&D intensity, gross profit ratio, and utilization-adjusted TFP growth rates for 1976-2020. In the second step, we pin down parameters by asking the model to reproduce the estimated trends in the data.

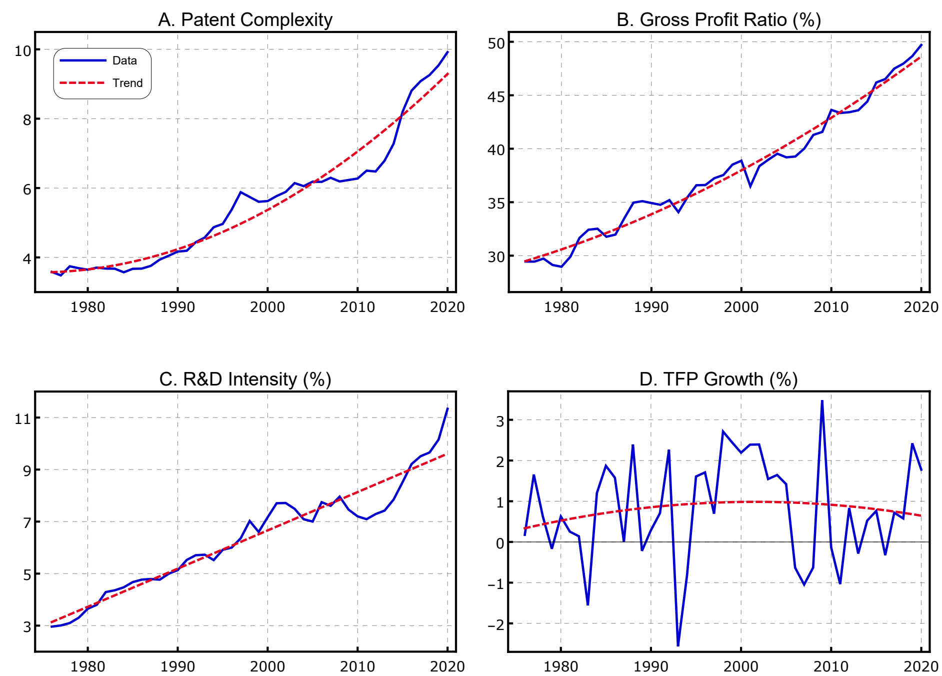
Trends in the Data Figure 2 overlays data and estimated trends for each variable. The trends are estimated by fitting quadratic polynomials to the raw data in the figure. As evident from the figure, patent complexity, gross profit ratios, and R&D intensity have all steadily risen since the mid 1970s.[10] By contrast, TPF growth rates are hump-shaped, with the peak occurring around 2000. As the comparative statics results in the previous section demonstrated, the model can reconcile the patterns of TFP growth, R&D intensity, and gross profit ratios with an increase in the scale of R&D operations, as implied by an increasing value of the overhead labor cost parameter, ϕ.

Matching the Trends in the Data There are two alternative yet complementary ways to quantify the model. While both are viable, they convey different lessons. One approach is to assume the economy is in a steady state at the start of the period, 1976, and at the end, 2020. Calibrate the model parameters to reproduce the initial steady state. Then, ask what value of the overhead cost parameter ϕ one would need to produce the values of R&D intensity, gross profit ratio, and TFP growth rate in 2020. This approach abstracts from the sizable time-series variation in observed TFP growth rates by construction. Had TFP growth been steadily increasing or decreasing throughout the period, such steady-state comparisons would capture the phenomenon's essence with little or no consequence on the lessons one draws. However, observed TFP growth rates are hump-shaped, rising from the early 1980s, plateauing from the late 1990s to the early 2000s, and declining since.

An alternative approach, which is the one we follow here, is to target the entire time series of R&D intensity, gross profit ratio, and TFP growth rates, requiring the model to fit

[10] In addition to the CPC count for patents granted to U.S. corporations, Appendix Figure *A*.4 shows rising CPC counts by all other major assignee types, including U.S. individuals, U.S. government, foreign corporations, foreign individuals, and foreign government.

Figure 2: Trends in the Data, 1976-2020



Notes: The figure shows raw data (solid line) and trends calculated based on quadratic polynomials (dashed line). Panel A shows the "patent complexity" measure constructed as the average count of cooperative patent classifications (CPC) in U.S. patent citations. Panel B shows the gross profit ratio, one minus COGS-to-Sales ratio, in high-R&D industries. Panel C shows the R&D-to-Sales ratio in high-R&D industries. Panel D shows utilization-adjusted TFP growth rates from Fernald (2014). See Appendix A for more details on data sources, variables' definitions, and construction.

them as closely as possible, akin to a moment-matching exercise, allowing the overhead labor cost parameter, ϕ_t, to change over time, while keeping all other parameters constant throughout the period. Here, too, there are two options. The first is to leave the path of the ϕ's completely unrestricted and back out the implied values by moment matching at the risk of making this approach overfitting. The second option is to discipline the path of the parameters with some extraneous information, which is what we do.

More specifically, we assume that ϕ_t depends linearly on patent complexity:

$$\phi_t = \varphi_0 + \varphi_1 \textit{Patent Complexity}_t + u_t, \tag{24}$$

where *Patent Complexity*$_t$ on the right-hand side is the patent complexity trend shown in panel A of Figure 2 and u_t is a regression residual. Operationally, this approach has the advantage of adding solely two parameters related to the intercept, φ_0, and the slope, φ_1, of the regression line (24). We then fix the time discount rate $\rho = 0.05$, and search over a grid of the remaining parameters including (φ_0, φ_1) to minimize the distance between the model transition paths and trends in the data.[11] This procedure gives the parameter values in Table 1.

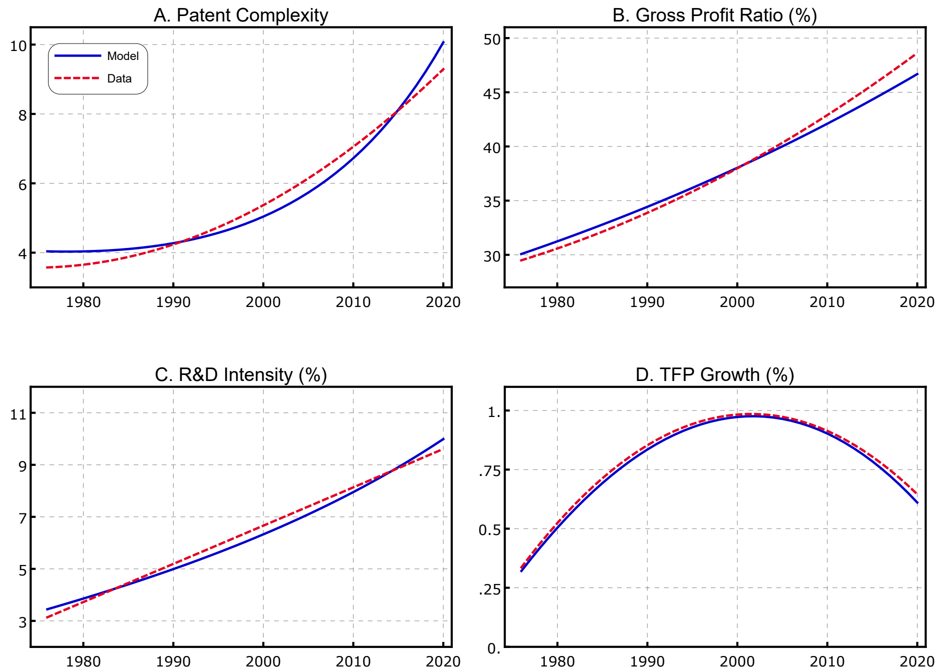
Table 1: Parameter Values

Parameter	Value	Parameter	Value
ϵ	8.50	θ	0.59
α	0.50	γ	6.66
δ	0.10	ρ	0.05
ω	4.00	ψ	0.45
φ_0	3.77	φ_1	1.69

As shown in Figure 3, the calibrated model fits remarkably well the time series of the variables of interest, including the hump-shaped path of TFP growth rates. Again, the driving force of the model's transition dynamics are the rising overhead labor costs of R&D labs. All the other parameters are fixed at their baseline values. Importantly, the hump-shaped TFP growth dynamics implied by the model is not hard-wired into the theory; rather, it is the result of the interplay of two opposing forces, pushing TFP growth in opposite direction as the market share of dominant firms rises, consistent with the data.

[11] We normalize population size to $L = 1$ throughout without loss of generality.

Figure 3: Model Fit, 1976-2020



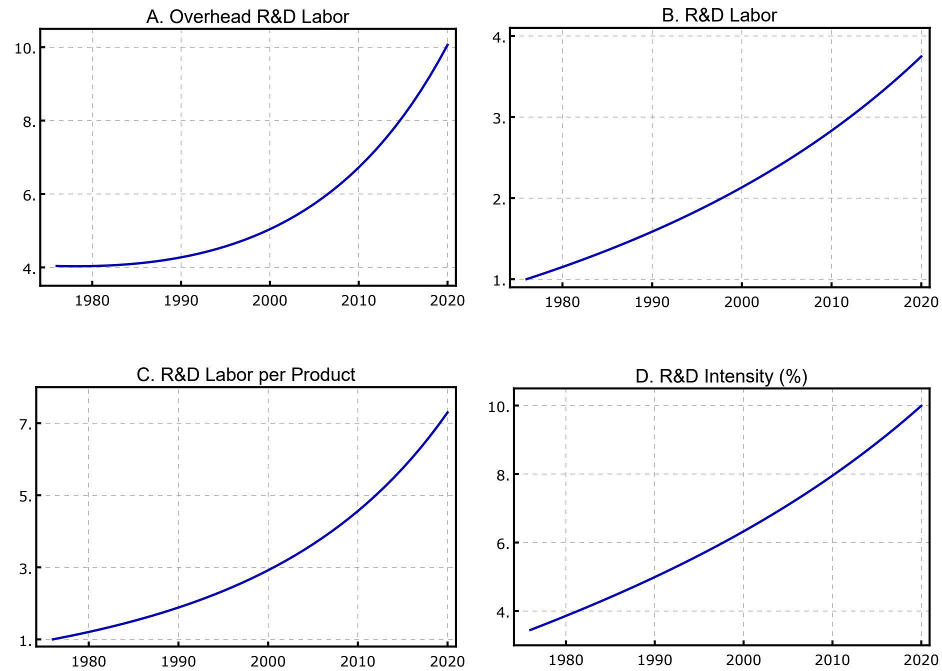
Notes: The figure shows model simulated time paths with perfect foresight (solid line) versus the trends in the data (dashed line) as shown in Figure 2. Panel A shows the time series of the overhead cost parameter ϕ_t for $t = 1976, \ldots, 2020$ as implied by the moment-matching exercise (solid line) and patent complexity (dashed line). Panel B shows the gross profit ratio (one minus the ratio of cost of goods sold (COGS) to sales) in the model calculated based on appendix equation (C.7) (solid line) and in the data (dashed line). Panel C shows R&D intensity (R&D-to-Sales ratio) in the model calculated based on appendix equation (C.8) (solid line) and in the data (dashed line). Panel D shows the TPF growth rate in the model calculated based on appendix equation (C.9) (solid line) and in the data (dashed line). See Appendix A for more details on data sources, variables' definitions, and construction.

5.3. Quantification

Panel A of Figure 4 shows the time path of the overhead R&D labor parameter implied by our moment-matching exercise. As evident from the figure, ϕ rises over time, mimicking the path of patent complexity in the data. Rising ϕ directly affects the optimal scale of R&D operations in the model as it effectively increases the minimum labor requirement for the firm's technology of knowledge creation. As a result, aggregate R&D labor and R&D labor per product increased from the mid-1970s onward; see panels B and C of the figure. Similarly, in panel D, R&D intensity steadily increases over time, consistent with the data. The model generates steadily increasing labor allocated to R&D at the aggregate, firm, or product level and as a share of firms' total sales.

Figure 5 shows that the model reconciles the rising R&D labor allocations and intensity

Figure 4: The Rising Scale of R&D Operations, 1976-2020



Notes: The figure shows model simulated time paths with perfect foresight. Panel A shows the path of the overhead R&D labor parameter (ϕ_t) as implied by the moment-matching exercise. Panels B, C, and D show the model simulated path for R&D labor in a dominant firm times the measure of dominant firms ($n\ell_z$), R&D labor in a dominant firm (ℓ_z), and R&D-intensity of a dominant firm (ℓ_z/px), respectively. In panels B and C, we normalize the values in 1976 to one.

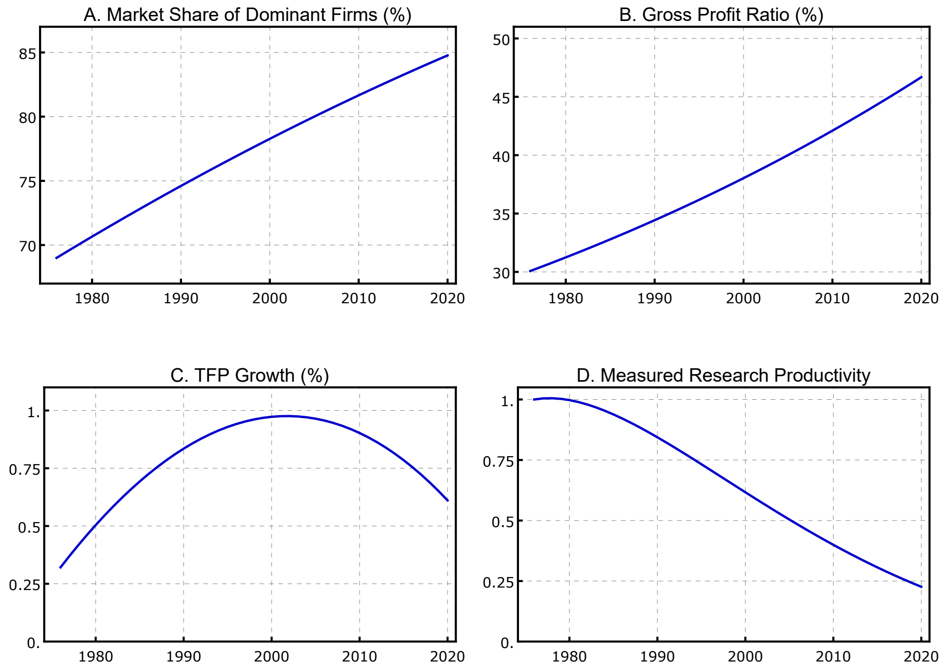
with the observed hump-shaped pattern of TFP growth while at the same time generating steadily rising market share of dominant firms and gross profit ratios.

Before describing how the current model generates these patterns, it is helpful to stress that a version of the model with constant markups would fail to do so. More specifically, a model with constant markups generates rising R&D labor insofar as the overhead labor parameter increases, yet it would grossly fall short in generating the observed TFP growth dynamics. In this sense, endogenous or variable markups are critical to our results.

In the model with variable markups, the equilibrium technology growth rate and how it changes over time depends on the relative strength of the cost-spreading and the business-stealing effects. The former implies a positive relationship between the market share of dominant firms and the growth rate; the latter implies a negative one.

First, the cost-spreading effect rests on a fundamental force at the core of invariably all endogenous growth theories and links the size of the market served by the firm to the firm's incentives to do R&D. As the market share of dominant firms grows large, an

Figure 5: Market Concentration and TFP Growth, 1976-2020



Notes: The figure shows model simulated time paths with perfect foresight after feeding to the model the path of ϕ_t in panel A of Figure 4. Panel A shows the market share of dominant firms (s_t). Panels B and C show the gross profit ratio and TFP growth rates calculated based on appendix equations (C.7) and (C.9), respectively. Panel D shows "measured research productivity," defined as the productivity growth of a dominant firm divided by its R&D labor ($\mathfrak{z}/\ell_z$), for which we normalize the value in 1976 to one.

increasing share of the market is served by the firms doing R&D. From the individual firm's perspective, knowledge creation is equivalent to reducing unit production costs. The higher the firm's scale of operation, the larger the gains from allocating an additional unit of labor to the R&D lab away from output production. Early on, when the market share of the dominant firm is small, this mechanism is strong enough to dominate the offsetting mechanism due to business stealing.

Second, the business-stealing effect stems from the dominant firms internalizing their impact on the price of the product line. The dominant firm's product demand depends on the price it charges relative to the average price level of the product line. As increased R&D lowers the price of the dominant firm, it also reduces the price of the product line; this adverse price effect reduces the gains of allocating labor to R&D. Such an effect strengthens as the market share of the dominant firm rises since the aggregate price index moves in lockstep with the dominant firm's price; it becomes exceedingly challenging to steal market share of other firms.[12]

Finally, panel D of Figure 5 shows that *measured* research productivity, defined as the ratio of the model's TFP growth rate to R&D labor, steadily declined throughout the post-1975 period. Such a downward trend materializes despite steadily increasing innovative investment as captured by ever-rising R&D labor at the firm and aggregate level and as a share of firms' sales. Again, these model's results do not come from changes in the productivity parameter that governs the knowledge production possibility frontier (α); instead, it is the outcome of rising overhead labor costs and endogenous marker structure.

6. Conclusion

This paper proposes a theory based on rising overhead costs of R&D to reconcile the hump-shaped pattern of TFP growth rates with the increasing R&D intensity in the United States from the mid-1970s onward. We quantify the theory in a new endogenous growth model and show that the model's predictions are borne out in the data.

The model features dominant firms competing with a fringe in the product market for imperfectly substitutable goods. Dominant firms charge variable markups and innovate by allocating labor to knowledge production, which reduces the firm's unit production costs. This firm-specific knowledge creation advances frontier technology, allowing for long-run growth. The competitive fringe taps into public knowledge, competing with the dominant firms and pricing at marginal cost. Aggregate productivity growth is hump-shaped in the market share of dominant firms, resulting from the offsetting effects of two opposing forces: cost-spreading versus business-stealing. The first dominates early when market shares are relatively small; the second effect dominates later when firms reach a critical market share. The calibrated model successfully reproduces the observations on TFP growth rates and R&D intensity, generating rising profits-to-sales ratios consistent with firm-level data from Compustat.

Here, we take the overhead cost parameters as model inputs. We view this approach as a first step towards theories of how R&D labs or innovative research teams more broadly work and their cost structure. Our quantitative results suggest that understanding the overhead costs of knowledge production and how they evolve can be a fruitful area of future research.

[12] Appendix Figure D.1 shows that the interest rate remains virtually constant along the transition path.

References

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li. 2023. "A Theory of Falling Growth and Rising Rents." *Review of Economic Studies*, 90(6): 2675–2702.

Aharonson, Barak S., and Melissa A. Schilling. 2016. "Mapping the Technological Landscape: Measuring Technology Distance, Technological Footprints, and Technology Evolution." *Research Policy*, 45(1): 81–96.

Benner, Mary, and Joel Waldfogel. 2008. "Close to You? Bias and Precision in Patent-Based Measures of Technological Proximity." *Research Policy*, 37(9): 1556–1567.

Bloom, Nicholas, Charles I. Jones, and John Van Reenen. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review*, 110(4): 1104–1144.

Cohen, Wesley M. 2010. "Fifty Years of Empirical Studies of Innovative Activity and Performance." *Handbook of the Economics of Innovation*, 1: 129–213.

Cohen, Wesley M., and Daniel A. Levinthal. 1989. "Innovation and Learning: The Two Faces of R&D." *Economic Journal*, 99(397): 569–596.

Cohen, Wesley M., and Richard C. Levin. 1989. "Empirical Studies of Innovation and Market Structure." *Handbook of Industrial Organization*, 2: 1059–1107.

Cowen, Tyler. 2011. *The Great Stagnation: How America Ate All the Low-hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better.* Penguin.

De Ridder, Maarten. 2024. "Market Power and Innovation in the Intangible Economy." *American Economic Review*, 114(1): 199–251.

Fernald, John. 2014. "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity." Federal Reserve Bank of San Francisco.

Ferraro, Domenico, Soroush Ghazi, and Pietro F Peretto. 2023. "Labour Taxes, Market Size and Productivity Growth." *Economic Journal*, uead028.

Fleming, Lee, and Olav Sorenson. 2001. "Technology as a Complex Adaptive System: Evidence from Patent Data." *Research Policy*, 30(7): 1019–1039.

Gomme, Paul, B. Ravikumar, and Peter Rupert. 2011. "The Return to Capital and the Business Cycle." *Review of Economic Dynamics*, 14(2): 262–278.

Gordon, Robert. 2017. *The Rise and Fall of American Growth: The US Standard of Living since the Civil War.* Princeton University Press.

Hall, Bronwyn H, Jacques Mairesse, and Pierre Mohnen. 2010. "Measuring the Returns to R&D." In *Handbook of the Economics of Innovation*. Vol. 2, 1033–1082. Elsevier.

Hopenhayn, Hugo, Julian Neira, and Rish Singhania. 2022. "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share." *Econometrica*, 90(4): 1879–1914.

Jones, Benjamin F. 2009. "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?" *Review of Economic Studies*, 76(1): 283–317.

Jones, Benjamin F. 2021. "The Rise of Research Teams: Benefits and Costs in Economics." *Journal of Economic Perspectives*, 35(2): 191–216.

Jones, Charles I., and John C. Williams. 2000. "Too Much of a Good Thing? The Economics of Investment in R&D." *Journal of Economic Growth*, 5(1): 65–85.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2019. "Demographic Origins of the Startup Deficit." *NBER Working Paper*, 25874.

Olmstead-Rumsey, Jane. 2022. "Market Concentration and the Productivity Slowdown." *Working Paper*.

Peretto, Pietro F. 1998. "Technological Change and Population Growth." *Journal of Economic Growth*, 3(4): 283–311.

Peretto, Pietro F., and Michelle Connolly. 2007. "The Manhattan Metaphor." *Journal of Economic Growth*, 12(4): 329–350.

Peretto, Pietro F., and Sjak Smulders. 2002. "Technological Distance, Growth and Scale Effects." *Economic Journal*, 112(481): 603–624.

Appendix

A. Data Sources and Additional Facts

A.1. Data Sources

Data sources are:

- Utilization-adjusted total factor productivity data from John Fernald's website are available at https://www.johnfernald.net/TFP.
- Financial data of the publicly traded firms in the U.S., used in the construction of COGS-to-Sales and R&D-to-Sales ratios, are from the Fundamentals Annual file of Compustat-CRSP merge database, available at https://wrds-www.wharton.upenn.edu.
- Data on cooperative patent classifications (CPCs) cited by patents granted to U.S. companies and corporations is constructed from the USPTO's PantentsView files, available at https://patentsview.org/download/data-download-tables.
- NAICS codes used to classify industries into high-R&D and low-R&D in Appendix Table A.1 are from the U.S. Census website, available at https://www.census.gov/naics/.

A.2. Additional Facts

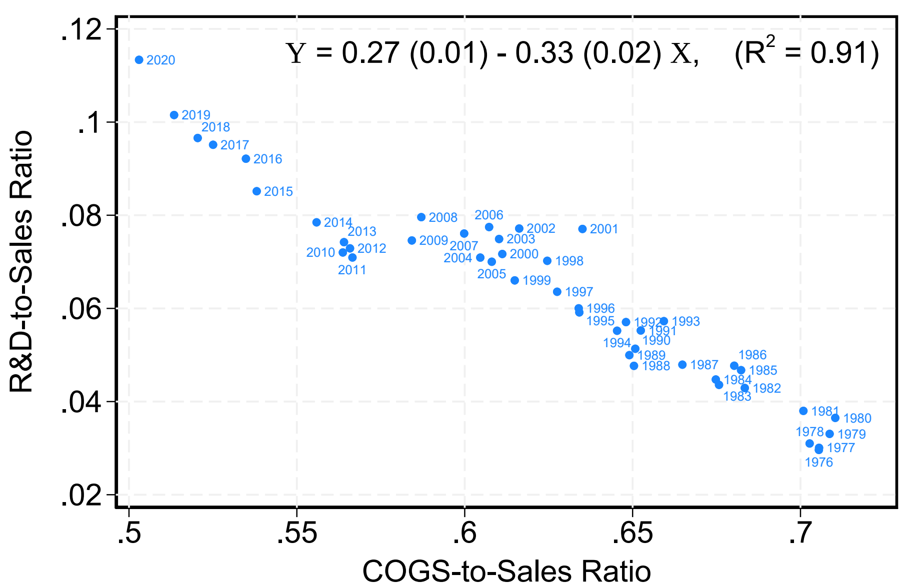
This subsection provides additional facts based on Compustat data.

A.2.1. Calculating Firm's Value Using Data on R&D-to-Sales and COGS-to-Sales

In the model, free entry implies that the cost of entry equals the post-entry value—the present discounted value of future dividends—of creating a new product. Here, we show how to calculate the model implied long-run value of firms based on the R&D-to-Sales and COGS-to-Sales ratios. Notably, the model predicts a linear relationship between these ratios as long as the firm's value is proportional to the gross profit flow, as implied by our specification of the free-entry condition. Data from Compustat strongly support this model prediction.

To see this, Figure $A.1$ shows R&D-to-Sales and COGS-to-Sales ratios along with the regression line and associated R-squared for our sample of top-20 high-R&D industries for 1976-2020. A dot represents a combination of annual averages of R&D-to-Sales and COGS-to-Sales ratios. The estimated slope coefficient $\hat{\beta}_1 = -0.33$ points to a negative relationship between the two ratios as predicted by the model. Note also that the estimate of the intercept $\hat{\beta}_0 = 0.27$ also supports the additional restriction $\hat{\beta}_0 \approx \hat{\beta}_1$, which we use below to simplify calculations.

Figure $A.1$: R&D-to-Sales vs. COGS-to-Sales, 1976-2020



Notes: The figure shows the R&D-to-Sales versus COGS-to-Sales ratios for the U.S. top-20 R&D industries in Compustat from 1976 to 2020. The estimated regression line and R^2 are reported as text in the figure.

Back to the model, we can rewrite a firm's value in the steady state as

$$v \approx \frac{1}{\rho + \delta}(\text{Sales} - \text{COGS} - \text{R\&D}), \tag{A.1}$$

assuming $r \approx \rho$ and so approximately constant over time and that the fixed operating cost of production is small relative to other costs. If we divide ($A.1$) by sales and use $\hat{\beta}_0 \approx \hat{\beta}_1$ so that $\frac{\text{R\&D}}{\text{Sales}} \approx \beta_0 - \beta_0 \frac{\text{COGS}}{\text{Sales}}$, we can rewrite ($A.1$) as

$$\frac{v}{\text{Sales}} \approx \frac{1}{\rho + \delta}\left(1 - \frac{\text{COGS}}{\text{Sales}} - \beta_0 + \beta_0 \frac{\text{COGS}}{\text{Sales}}\right) = \frac{1 - \beta_0}{\rho + \delta}\left(1 - \frac{\text{COGS}}{\text{Sales}}\right). \tag{A.2}$$

Finally, note that the COGS-to-Sales ratio in the model equals the inverse markup $\left(\frac{e-1}{e}\right)$

so that

$$\frac{v}{\text{Sales}} \approx \left(\frac{1-\beta_0}{\rho+\delta}\right)\frac{1}{e}. \tag{A.3}$$

Thus, the model-implied long-run value of the firm is proportional to the gross profit flow (px/e) to the extent that the R&D-to-Sales ratio is approximately proportional to the COGS-to-Sales ratio. To be sure, this argument hinges on the premise that the scatter plot above describes reasonably well the long-run relationship between COGS-to-Sales and R&D-to-Sales in the data and that the fixed costs of output production are small relative to COGS and R&D, which would seem plausible for high-R&D industries.

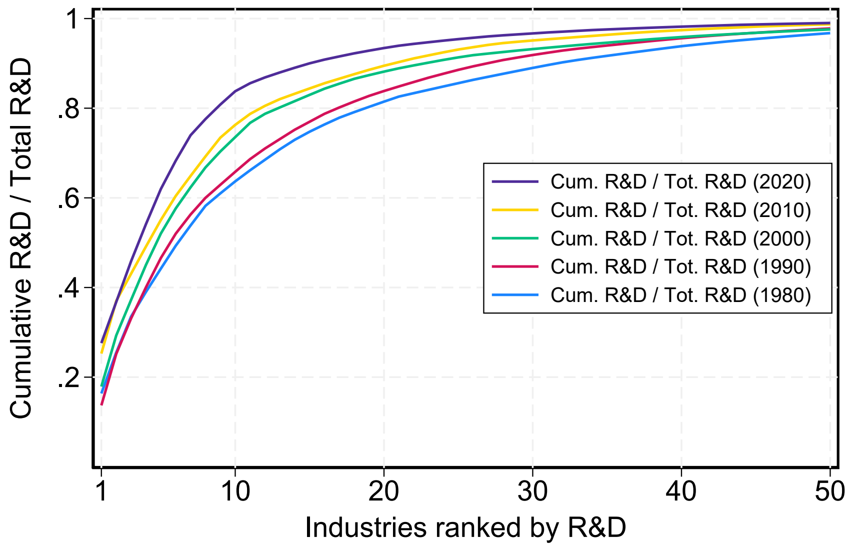
A.2.2. R&D by Industry

This subsection details that a few industries in Compustat account for most of R&D in the data and that there is a meaningful distinction between large firms measured in terms of their employment and high-R&D firms. While there is naturally some overlap between the two categories, they do not entirely coincide.

High-R&D Industries We rank 4-digit NAICS industries in Compustat by their R&D each year. For each rank r, we then calculate the share of R&D that is performed by industries ranked up to r. Figure *A.2* shows R&D is highly concentrated in a few top-ranked industries. For instance, in 2020, the top ten industries performed about 84% and the top twenty industries performed more than 93% of total R&D. Figure *A.2* also shows that cumulative R&D curves have gradually shifted since 1980, revealing that the concentration of R&D in the top industries has been rising decade after decade.

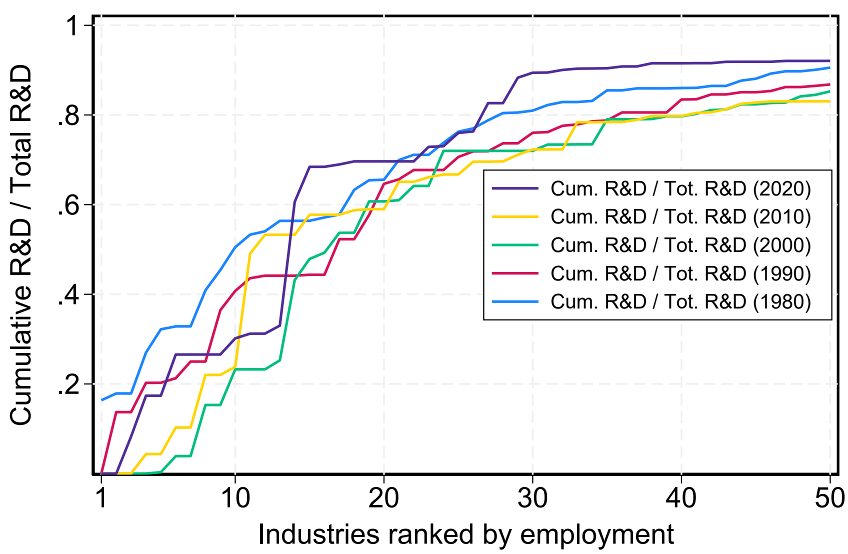
Table *A.1* lists the NAICS titles for the sample of the top twenty high-R&D industries in Section 5. These are industries with the highest aggregate (real) R&D expenditure in Compustat data for 2017-2021. As evident from the table, the high-R&D industries are mainly in manufacturing, with the pharmaceutical and semiconductor industries being the largest. The high-R&D service industries in the sample are web services and software publishers. NAICS 9999 denotes other industries, which we excluded from our analysis since it potentially contains firms from many heterogeneous industries. Keeping them in the sample makes little difference for the results.

Large Firms vs. High-R&D Firms As is well-documented, larger industries spend on average more on R&D (Cohen, 2010; Cohen and Levin, 1989). Here, we highlight that while there is some overlap between large firms by employment versus high-R&D firms, Figure $A.2$: Cumulative R&D by R&D-Ranked Industries



Notes: Industries are ranked by their R&D expenditures, e.g., in 2020, the industry ranked 1 has the highest R&D, the industry ranked 2 has the second highest R&D, and so on. For rank r on the x-axis, the figure shows the fraction of total R&D by industries ranked 1 up to r on the y-axis.

Figure $A.3$: Cumulative R&D by Employment-Ranked Industries



Notes: Industries are ranked by their employment, e.g., in 2020, the industry ranked 1 has the highest employment, the industry ranked 2 has the second highest employment, and so on. For rank r on the x-axis, the figure shows the fraction of total R&D by industries ranked 1 up to r on the y-axis.

the two classifications do not coincide. Indeed, ranking industries by size rather than R&D provides a rather different picture of R&D by industry. Figure $A.3$ shows cumulative R&D by industries ranked by their employment as opposed to R&D. The figure shows the fraction of R&D expenditure performed by employment-ranked industries from rank 1 up to r. In 2020, only 30% of total R&D is performed by the top ten largest industries and 70% of total R&D is performed by the top twenty largest industries. R&D is considerably less concentrated by firm size.

Table A.1: High-R&D Industries

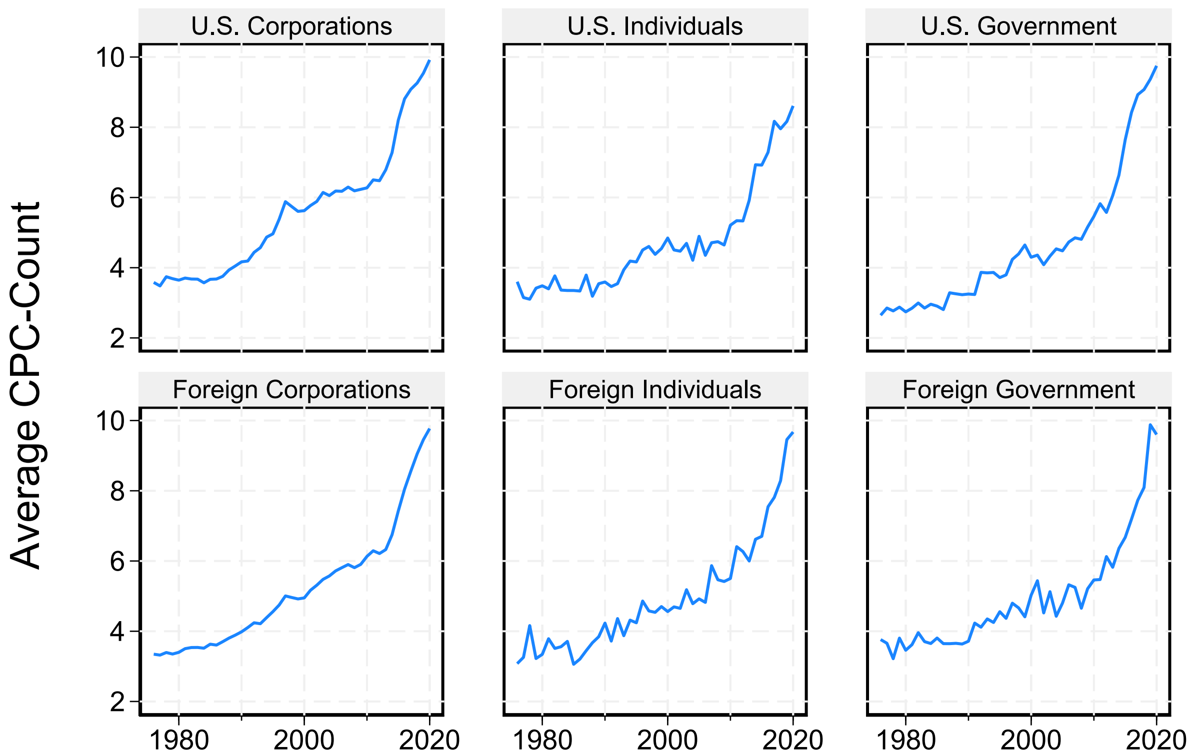
Rank	NAICS	NAICS Title
1	3254	Pharmaceutical and Medicine Manufacturing
2	3344	Semiconductor and Other Electronic Component Manufacturing
3	5182	Computing Infrastructure Providers, Data Processing, Web Hosting, and Related Services
4	4541*	Electronic Shopping and Mail-Order Houses
5	5192	Web Search Portals, Libraries, Archives, and Other Information Services
6	5132	Software Publishers
7	3342	Communications Equipment Manufacturing
8	3361	Motor Vehicle Manufacturing
9	3345	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing
10	3341	Computer and Peripheral Equipment Manufacturing
11	3364	Aerospace Product and Parts Manufacturing
12	5112*	Software Publishers
13	3391	Medical Equipment and Supplies Manufacturing
14	5191*	Other Information Services
15	5415	Computer Systems Design and Related Services
16	3332	Industrial Machinery Manufacturing
17	3331	Agriculture, Construction, and Mining Machinery Manufacturing
18	3363	Motor Vehicle Parts Manufacturing
19	9999	Other Industries in Compustat
20	3256	Soap, Cleaning Compound, and Toilet Preparation Manufacturing

Notes: The table shows the NAICS titles of high-R&D industries in Compustat for 2017-2021. Titles are based on 2022 NAICS definitions from the U.S. Census, except for NAICS marked with *, which are 2017 NAICS and are discontinued in 2022 NAICS. The reason is that many firms in our database are assigned to older classifications. NAICS 9999 represents firms listed as other NAICS in Compustat.

A.2.3. Cooperative Patent Classification (CPC) Counts by Assignee Type

This subsection provides additional facts on the cooperative patent classification (CPC) counts by assignee type over time. Recall that Figure 2 in the paper's main text considers solely U.S. corporations and companies. Figure A.4 below reports the CPC count for other major assignee types, including corporations, individuals, and government, both U.S. and foreign. The figure shows a similar increasing trend in CPC counts across all assignee types.

Figure $A.4$: Average CPC Count by Assignee Type



Notes: CPC-count is the number of cooperative patent classifications to which a patent belongs, which is a measure or proxy for "patent complexity." The figure shows the annual average CPC count for major types of patent assignee. Data source is USTPO PatentsView files for 1976-2020.

B. Derivations and Proofs

This appendix details the derivations of the equations in the paper's main text and proofs.

B.1. Derivations of Model Equations

Derivation of Equation (13) Since $W = 1$, the dividend flow is $d_i = p_i x_i - (\ell_{x,i} + \ell_{z,i})$. Replacing $\ell_{x,i} = \frac{x_i}{z_i^\theta Z^{1-\theta}} + \psi$ from (8), and $x_i = Y_i \frac{p_i^{-\epsilon}}{P_i^{1-\epsilon}}$ from (5), we obtain

$$d_i = Y_i \frac{p_i^{1-\epsilon}}{P_i^{1-\epsilon}} - \frac{1}{z_i^\theta Z^{1-\theta}} Y_i \frac{p_i^{-\epsilon}}{P_i^{1-\epsilon}} - \psi - \ell_{z,i}. \tag{B.1}$$

Hence, the current value Hamiltonian is

$$\mathcal{H}^c = Y_i \frac{p_i^{1-\epsilon}}{P_i^{1-\epsilon}} - \frac{1}{z_i^\theta Z^{1-\theta}} Y_i \frac{p_i^{-\epsilon}}{P_i^{1-\epsilon}} - \psi - \ell_{z,i} + \lambda\, \alpha Z(\ell_{z,i} - \phi), \tag{B.2}$$

where λ is the costate variable. The first-order conditions are

$$\frac{\partial \mathcal{H}^c}{\partial \ell_{z,i}} = 0, \tag{B.3}$$

$$(r+\delta)\lambda = \dot{\lambda} + \frac{\partial \mathcal{H}^c}{\partial z_i}. \tag{B.4}$$

The first equation above yields the shadow value of technology $\lambda = 1/(\alpha Z)$, and the second equation gives

$$(r+\delta)\lambda = \dot{\lambda} + \frac{\theta}{z_i} \frac{1}{z_i^\theta Z^{1-\theta}} Y_i \frac{p_i^{-\epsilon}}{P_i^{1-\epsilon}}. \tag{B.5}$$

Dividing through by λ and using $\lambda = 1/(\alpha Z)$ gives

$$r + \delta = -\frac{\dot{Z}}{Z} + \alpha\theta \frac{Z}{z_i} \frac{1}{z_i^\theta Z^{1-\theta}} \frac{1}{p_i} Y_i \frac{p_i^{1-\epsilon}}{P_i^{1-\epsilon}}, \tag{B.6}$$

where we divided and multiplied the last term by p_i. Finally, using the markup price $p_i = \frac{e_i}{e_i - 1}\frac{1}{z_i^{\theta} Z^{1-\theta}}$ in (12) to replace $\frac{1}{z_i^{\theta} Z^{1-\theta}}\frac{1}{p_i}$ with $\frac{e_i - 1}{e_i}$, and using $s_i = \frac{p_i^{1-\epsilon}}{P_i^{1-\epsilon}}$ in (7), we obtain

$$r = -\frac{\dot{Z}}{Z} - \delta + \alpha\theta\frac{Z}{z_i} \cdot Y_i s_i \cdot \frac{e_i - 1}{e_i}. \tag{B.7}$$

Derivation of Equation (23) Start with no-arbitrage in symmetric equilibrium, $r + \delta = \frac{\dot{v} + d}{v}$. Recall that $v = \gamma\frac{sY}{e}$, or $v = \gamma\frac{s\mathcal{Y}}{enL}$. Replace v and $r = \rho + \frac{\dot{\mathcal{Y}}}{\mathcal{Y}}$, to obtain

$$\rho + \frac{\dot{\mathcal{Y}}}{\mathcal{Y}} + \delta = \frac{\dot{s}}{s} + \frac{\dot{\mathcal{Y}}}{\mathcal{Y}} - \frac{\dot{e}}{e} - \frac{\dot{n}}{n} + \frac{d}{(\gamma\frac{s\mathcal{Y}}{enL})}, \tag{B.8}$$

where the $\frac{\dot{\mathcal{Y}}}{\mathcal{Y}}$ drops from both sides. Next, use $d = \frac{s\mathcal{Y}}{enL} - \frac{\mathfrak{z}}{\alpha n} - (\psi + \phi)$ and $\mathfrak{z} = -(r + \delta) - \frac{\dot{n}}{n} + \alpha\theta\frac{\mathcal{Y}}{L}s\frac{e-1}{e}$ from (22), and rearrange, to obtain

$$\frac{\dot{n}}{n}\left(1 - \frac{eL}{\alpha\gamma s\mathcal{Y}}\right) + \frac{\dot{e}}{e} - \frac{\dot{s}}{s} = \frac{1 - \theta(e-1)}{\gamma} + \left(r + \delta\right)\frac{eL}{\alpha\gamma s\mathcal{Y}} - \frac{enL}{\gamma s\mathcal{Y}}(\psi + \phi), \tag{B.9}$$

which is (23) in the paper's main text.

Here, we rewrite the equation above solely in terms of s to show explicitly the second-order nonlinear differential equation in the dominant firm's market share, s. First, we use the expression for $\mathcal{Y}$ in (20), which we reproduce below for convenience,

$$\mathcal{Y} = \frac{L}{1 - \gamma\frac{s}{e}(\rho + \frac{\dot{e}}{e} - \frac{\dot{s}}{s})}, \tag{B.10}$$

and $r = \frac{\dot{\mathcal{Y}}}{\mathcal{Y}} + \rho$ to r as a function of s containing its first-order and second-order derivative, $\dot{s}$ and $\ddot{s}$. In the calculation, we use $\frac{d}{dt}(\frac{\dot{s}}{s}) = \frac{\ddot{s}}{s} - (\frac{\dot{s}}{s})^2$, and $\frac{d}{dt}(\frac{\dot{e}}{e}) = \frac{\ddot{e}}{e} - (\frac{\dot{e}}{e})^2$. Thus,

$$r = \frac{\dot{\mathcal{Y}}}{\mathcal{Y}} + \rho = \frac{(\frac{\dot{s}}{s} - \frac{\dot{e}}{e})(\rho + \frac{\dot{e}}{e} - \frac{\dot{s}}{s}) + \left[\frac{\ddot{e}}{e} - (\frac{\dot{e}}{e})^2 - \frac{\ddot{s}}{s} + (\frac{\dot{s}}{s})^2\right]}{1 - \gamma\frac{s}{e}(\rho + \frac{\dot{e}}{e} - \frac{\dot{s}}{s})} \cdot \gamma\frac{s}{e} + \rho. \tag{B.11}$$

Finally, replacing n from (21) in (B.9), we rewrite (B.9) as the following second-order nonlinear differential equation:

$$
\begin{aligned}
&\frac{\epsilon}{\theta e(e-1)}\left\{1-\frac{e}{\alpha\gamma s}\left[1-\gamma\frac{s}{e}\left(\rho+\frac{\dot{e}}{e}-\frac{\dot{s}}{s}\right)\right]\right\}\frac{\dot{s}}{s}-\frac{\dot{e}}{e}+\frac{\dot{s}}{s}+\frac{1}{\gamma}\left[1-\theta(e-1)\right]\\
&+\left\{\frac{(\frac{\dot{s}}{s}-\frac{\dot{e}}{e})(\rho+\frac{\dot{e}}{e}-\frac{\dot{s}}{s})+\left[\frac{\ddot{e}}{e}-(\frac{\dot{e}}{e})^2-\frac{\ddot{s}}{s}+(\frac{\dot{s}}{s})^2\right]}{1-\gamma\frac{s}{e}(\rho+\frac{\dot{e}}{e}-\frac{\dot{s}}{s})}\cdot\gamma\frac{s}{e}+\rho+\delta\right\}\\
&\times\frac{e}{\alpha\gamma s}\left[1-\gamma\frac{s}{e}\left(\rho+\frac{\dot{e}}{e}-\frac{\dot{s}}{s}\right)\right]\\
&-\frac{e}{\gamma s}\left[1-\gamma\frac{s}{e}\left(\rho+\frac{\dot{e}}{e}-\frac{\dot{s}}{s}\right)\right]\omega^{\frac{-\epsilon}{(\epsilon-1)\theta}}\left(\frac{e-1}{e}\right)^{\frac{1}{\theta}}\left(\frac{1-s}{s}\right)^{\frac{1}{(\epsilon-1)\theta}}(\psi+\phi)=0,
\end{aligned}
\tag{B.12}
$$

where $e=\epsilon-(\epsilon-1)s$.

Derivation of Equation (C.7) In the model, labor is the only factor of production, and the gross profit flow is $px-(l_x-\psi)$, in which $\ell_x-\psi$ is the cost of goods sold, COGS, and ψ is the fixed operating cost of production. Note that given the dominant firms' production technology in (8), in the symmetric equilibrium, we have $x=z^{\theta}Z^{1-\theta}(\ell_x-\psi)$, and hence, we have the marginal cost $mc=1/(z^{\theta}Z^{1-\theta})$. Thus, COGS $=\ell_x-\psi=x\times mc=xp\times\frac{mc}{p}=px\frac{e-1}{e}$, where $\frac{mc}{p}$ is the inverse markup. In short, COGS $=$ Sales $\times\frac{e-1}{e}$, and we have $\frac{\text{COGS}}{\text{Sales}}=\frac{e-1}{e}$, that is, COGS-to-Sales ratio equals the inverse markup. Thus,

$$
\text{gross profit ratio}=1-\frac{\text{COGS}}{\text{Sales}}=\frac{1}{e},
\tag{B.13}
$$

that is, the gross profit ratio equals the inverse of the price elasticity of demand.

Derivation of Equation (C.8) R&D labor is denoted by ℓ_z. Here, we are looking for the ratio $\frac{\ell_z}{px}$. Start from the no-arbitrage equation $r+\delta=\frac{\dot{v}+d}{v}$, and replace the dominant firm's value $v=\gamma\frac{s\mathcal{Y}}{enL}$, and $r=\rho+\frac{\dot{\mathcal{Y}}}{\mathcal{Y}}$, and drop $\frac{\dot{\mathcal{Y}}}{\mathcal{Y}}$ from both sides to obtain

$$
\rho+\delta=\frac{\dot{s}}{s}-\frac{\dot{e}}{e}-\frac{\dot{n}}{n}+\frac{d}{v}.
\tag{B.14}
$$

From the derivation of (C.7) above, we know that $\ell_x-\psi=px\frac{e-1}{e}$. Thus, one can write $d=px-\ell_x-\ell_z=px-(\ell_x-\psi)-(\psi+\ell_z)=\frac{px}{e}-(\psi+\ell_z)$. Also, note that $px=sY=$

$s\frac{\mathcal{Y}}{N} = \frac{s\mathcal{Y}}{nL}$, and $v = \gamma\frac{px}{e}$. Substituting for d and v in the above equation, we obtain

$$\rho + \delta = \frac{\dot{s}}{s} - \frac{\dot{e}}{e} - \frac{\dot{n}}{n} + \frac{\frac{px}{e} - (\psi + \ell_z)}{\gamma\frac{px}{e}}, \tag{B.15}$$

which after rearrangement becomes

$$\frac{\ell_z}{px} \cdot \frac{e}{\gamma} = \frac{1}{\gamma} - (\rho + \delta) + \frac{\dot{s}}{s} - \frac{\dot{e}}{e} - \frac{\dot{n}}{n} - \frac{\psi}{\gamma\frac{px}{e}}. \tag{B.16}$$

Finally, multiplying both sides by γ/e and replacing $px = \frac{s\mathcal{Y}}{nL}$, yields

$$\frac{\ell_z}{px} = \frac{1}{e}\left[1 - (\rho + \delta)\gamma + \left(\frac{\dot{s}}{s} - \frac{\dot{e}}{e} - \frac{\dot{n}}{n}\right)\gamma\right] - \frac{\psi nL}{s\mathcal{Y}}, \tag{B.17}$$

which is (C.8).

Derivation of Equation (C.9) As there is no "love of variety" in the specification of the final good production function (1), it is suffice to calculate the growth rate of technology of a product line, which consists of a dominant firm whose market share is s and its competitive fringe whose market share is $1 - s$. Thus, the aggregate (measured) TFP growth rate equals the growth rate of productivity of the dominant firm times s, plus the growth rate of productivity of the competitive fringe times $1 - s$.

From (8), the productivity a dominant firm is $z^\theta Z^{1-\theta} = z^\theta (nz)^{1-\theta} = zn^{1-\theta}$ and its TFP growth rate is $\mathfrak{z} + (1-\theta)\frac{\dot{n}}{n}$. From (14), the productivity of firms in the fringe is $Z = nz$ and its TFP growth rate is $\mathfrak{z} + \frac{\dot{n}}{n}$. Finally, the measured TFP growth of the economy is

$$\text{TFP growth rate} = s\left[\mathfrak{z} + (1-\theta)\frac{\dot{n}}{n}\right] + (1-s)\left(\mathfrak{z} + \frac{\dot{n}}{n}\right) = \mathfrak{z} + (1 - \theta s)\frac{\dot{n}}{n}, \tag{B.18}$$

which is (C.9).

B.2. Proofs

Proof of Proposition 1 We show that for a dominant firm i that starts with technology level z_i and takes everything in the economy except its price and investment as given, the return on R&D is decreasing in z_i. This condition guarantees that a firm that starts with a smaller (larger) than average z_i grows faster (slower) than the average dominant firm, and hence, its technology level should converge to the average private technology level in the economy. As a result, on the balanced growth path, all firms converge to the average.

Start with the rate of return to dominant firm i's R&D in (13),

$$r = -\delta - \frac{\dot{Z}}{Z} + \alpha\theta \frac{Z}{z_i} \cdot Y_i s_i \cdot \frac{e_i - 1}{e_i}. \tag{B.19}$$

Notably, the technology ratio $\frac{Z}{z_i}$ in this economy determines the market shares s_i, and there is a one-to-one map between the two (i.e., a higher technology z_i is equivalent to a higher market share s_i). Below, we establish this property, rewrite the return solely in terms of s_i, and find the condition that guarantees the return is decreasing in s_i, which we show to be $\underline{s} < s_i$.

To find the map between s_i and $\frac{Z}{z_i}$, start with the dominant firm's market share in (7),

$$\begin{aligned} s_i = \frac{p_i^{1-\epsilon}}{p_i^{1-\epsilon} + \omega^\epsilon p_{0i}^{1-\epsilon}} = \frac{1}{1 + \omega^\epsilon \left(\frac{p_i}{p_{0i}}\right)^{1-\epsilon}} &= \frac{1}{1 + \omega^\epsilon \left(\frac{e_i}{e_i - 1} \cdot \frac{Z}{z_i^\theta Z^{1-\theta}}\right)^{1-\epsilon}} \\ &= \frac{1}{1 + \omega^\epsilon \left[\frac{e_i}{e_i-1} \left(\frac{Z}{z_i}\right)^\theta\right]^{1-\epsilon}}, \end{aligned} \tag{B.20}$$

and solve the above for $\frac{Z}{z_i}$ to obtain

$$\frac{Z}{z_i} = \omega^{\frac{-\epsilon}{(\epsilon-1)\theta}} \left(\frac{1 - s_i}{s_i}\right)^{\frac{1}{(\epsilon-1)\theta}} \left[\frac{(\epsilon - 1)(s_i - 1)}{\epsilon - (\epsilon - 1)s_i}\right]^{\frac{1}{\theta}}, \tag{B.21}$$

where we also substituted $e_i = \epsilon - (\epsilon - 1)s_i$. The partial derivative of the right-hand side with respect to s_i is

$$-\frac{\omega^{\frac{-\epsilon}{(\epsilon-1)\theta}}}{\theta s_i e_i^2} \left(\frac{1 - s_i}{s_i}\right)^{\frac{1}{(\epsilon-1)\theta}} \left(\frac{e_i - 1}{e_i}\right)^{\frac{1}{\theta} - 1} < 0. \tag{B.22}$$

Thus, $\frac{Z}{z_i}$ is strictly decreasing in s_i so that z_i is strictly increasing in s_i.

Having shown that z_i and s_i move in the same direction, what remains to show is to find the condition that guarantees the return to R&D is decreasing in s_i. Rewriting the return equation above only in terms of s_i, yields

$$r = -\delta - \frac{\dot{Z}}{Z} + \alpha\theta\omega^{\frac{-\epsilon}{(\epsilon-1)\theta}} \frac{Y}{N} s_i \left(\frac{1-s_i}{s_i}\right)^{\frac{1}{(\epsilon-1)\theta}} \left[\frac{(\epsilon-1)(s_i-1)}{\epsilon-(\epsilon-1)s_i}\right]^{\frac{1}{\theta}}. \tag{B.23}$$

Taking the partial derivative with respect to s_i and setting it equal to zero gives one positive solution less than one, i.e., $\underline{s} = \frac{2\epsilon-1-\sqrt{1+4\epsilon/\theta}}{2(\epsilon-1)}$, which is also a maximum. The other root of the derivative is at the corner $s_i = 1$, and there is no root in between. Thus, the return decreases for any $s_i \in [\underline{s}, 1]$. Finally, if a dominant firm starts with technology z_i so small that $s_i < \underline{s}$, then that firm never catches up with the other growing dominant firms. Thus, its market share approaches zero, eventually dropping out of the economy. □

Proof of Corollary 2 The expression of $n(s)$ from (21) is reproduced below, where we substituted $e = \epsilon - (\epsilon - 1)s$,

$$n(s) = \omega^{\frac{-\epsilon}{(\epsilon-1)\theta}} \left[\frac{(\epsilon-1)(s-1)}{\epsilon-(\epsilon-1)s}\right]^{\frac{1}{\theta}} \left(\frac{1-s}{s}\right)^{\frac{1}{(\epsilon-1)\theta}}. \tag{B.24}$$

The derivative of $n(s)$ with respect to s, which is negative for $s \in (0,1)$, is

$$n'(s) = -\frac{\omega^{\frac{-\epsilon}{(\epsilon-1)\theta}}}{\theta s\left[\epsilon-(\epsilon-1)s\right]^2} \left(\frac{1-s}{s}\right)^{\frac{1}{(\epsilon-1)\theta}} \left[\frac{(\epsilon-1)(s-1)}{\epsilon-(\epsilon-1)s}\right]^{\frac{1}{\theta}-1} < 0, \text{ if } s \in (0,1). \tag{B.25}$$

Note also that the derivative approaches $-\infty$ if s approaches zero, and it approaches zero if s approaches one. □

Proof of Proposition 3 There are no exogenous changes on the balanced growth path (BGP), i.e., ϕ is fixed, and endogenous variables either don't change or grow at a constant rate. In particular, we have $\dot{s} = 0$. Thus, the no-arbitrage condition $r = (\dot{v} + d)/v - \delta$, combined with (19) implies $\rho + \delta = d/v$. Substituting v and d yields $\frac{d}{v} = (\frac{s\mathcal{Y}}{eN} - \psi - \frac{\mathfrak{z}}{\alpha n} - \phi)\frac{eN}{\gamma s\mathcal{Y}}$. And substituting $\mathfrak{z}$ using (22) and simplifying, gives

$$\frac{d}{v} = \frac{1-\theta(e-1)}{\gamma} + (\rho+\delta)\frac{eL}{\alpha\gamma s\mathcal{Y}} - \frac{eN}{\gamma s\mathcal{Y}}(\phi+\psi). \tag{B.26}$$

Thus, we can rewrite $\rho + \delta = d/v$ as

$$\frac{1-\theta(e-1)}{\gamma} + (\rho+\delta)\Big(\frac{eL}{\alpha\gamma s\mathcal{Y}} - 1\Big) - \frac{eN}{\gamma s\mathcal{Y}}(\phi+\psi) = 0. \tag{B.27}$$

To find the effects of ϕ, ψ, and α on s (on the BGP), rewrite the above as

$$\big[1-\theta(e-1)\big]\frac{s\mathcal{Y}}{eN} + (\rho+\delta)\frac{1}{\alpha n} = \phi + \psi + (\rho+\delta)\frac{\gamma s\mathcal{Y}}{eN}, \tag{B.28}$$

in which both sides of the equation are increasing in s (for the first part on the left-hand side, we also need θ not too large). For instance, to see that, $\frac{s\mathcal{Y}}{eN}$ is increasing in s, note that both $e = \epsilon - (\epsilon - 1)s$ and $N = Ln$ are decreasing in s (to see the latter, use $N = nL$, (21), and Corollary 2). Moreover, from (20), on the BGP we have $\mathcal{Y} = \frac{L}{1-\gamma\rho\frac{s}{e}}$, which is also increasing in s. Put together, we conclude that $\frac{s\mathcal{Y}}{eN}$ is increasing in s, implying that the gross profit flow of dominant firms is increasing in their market shares. Finally, note that none of the variables $\mathcal{Y}$, e, and n in the equation above explicitly depend on either of parameters ϕ, ψ, and α.

First, note that as either ϕ or ψ increases, the right-hand side shifts upward; hence, s increases if the slope of the right-hand side is flatter than the left-hand side at their crossing point, which is obtained if γ is not too large. Second, as α decreases, the left-hand side shifts upward; hence, s decreases if the slope of the right-hand side is flatter than the left-hand side at their crossing point, which is obtained if γ is not too large.

As a side note, the slope of both sides for large values of s (that are empirically relevant for us) can be approximated by the $1/n$ since it is evident from the expression of $n'(s)$ in (B.22) that the slope of $1/n$ goes to infinity as s approaches one. Thus, the condition of θ not being too large is not numerically important in our application. □

Proof of Proposition 4 Part (1): The proof has two steps. First, on the balanced growth path, s increases in ϕ and ψ, which we established in Proposition 3. Second, we can find the balanced growth path of $\mathfrak{z}$ as a function of s, which has an inverted-U shape in s (all parameters but either ϕ or ψ are kept constant). Given these two steps, as ϕ increases from a small number, the balanced growth path s increases, which makes $\mathfrak{z}$ initially increase up to a point and then decline.

To see that $\mathfrak{z}$ on the balanced growth path (BGP) is hump-shaped in s when ϕ or ψ

changes, start with (22):

$$\mathfrak{z} = -(\rho+\delta) + \alpha\theta\frac{\mathcal{Y}}{L}\cdot s\cdot\frac{e-1}{e}, \tag{B.29}$$

where we set $\frac{\dot{n}}{n} = 0$ and $r = \rho$ since we are concerned with the BGP. The right-hand side in terms of s is

$$\mathfrak{z} = -(\rho+\delta) + \alpha\theta\frac{\epsilon-1}{\epsilon-(\epsilon+\rho\gamma-1)s}(1-s)s. \tag{B.30}$$

Notably, the right-hand side does not depend on either ϕ or ψ; hence, those parameters affect $\mathfrak{z}$ only through s. The second term on the right-hand side equals zero only at the two extremes $s = 0$ and $s = 1$. Moreover, assuming $\gamma < 1/\rho$ (which easily holds in our calibration), the derivative of the above equals zero only at the following point in the unit interval,

$$s_o = \frac{\epsilon - \sqrt{\epsilon(1-\rho\gamma)}}{\epsilon+\rho\gamma-1}, \tag{B.31}$$

which is a maximum since the second term is positive on the interval $s \in (0,1)$.

Part (2): This part has tow steps. First, we must show that s increases in α, which is done in Proposition 3. Second, we show that $n\ell_z$ does not directly depend on α and only indirectly increases in α through the rise of s. Then, we can use (10) on the BGP to conclude that $\mathfrak{z}$ increases in α.

To see that ℓ_z increases in s, start with the no-arbitrage condition $(\rho+\delta)v = d$ on the BGP. As in the proof of Proposition 3, replace $d = \frac{s\mathcal{Y}}{eN} - \psi - \ell_z$ and $v = \gamma\frac{s\mathcal{Y}}{eN}$. Solve for ℓ_z to obtain

$$\ell_z = \left[1-(\rho+\delta)\gamma\right]\frac{s\mathcal{Y}}{eN} - \psi. \tag{B.32}$$

Next, use $N = nL$ to rewrite the above as

$$n\ell_z = \left[1-(\rho+\delta)\gamma\right]\frac{s\mathcal{Y}}{eL} - \psi n = \left[1-(\rho+\delta)\gamma\right]\frac{s}{\epsilon-(\epsilon+\rho\gamma-1)s} - \psi n, \tag{B.33}$$

where the second equality follows since $e = \epsilon-(\epsilon-1)s$, and on the BGP $\mathcal{Y} = \frac{L}{1-\gamma\rho\frac{s}{e}}$. The expression for $n(s)$ and the fact that n decreases in s can be found in (21) and Corollary 2, respectively. Thus, the above expression increases in s, and importantly, it does not explicitly depend on α. Thus, α affects $n\ell_z$ only through s, that is, as α increases, so does s, and hence, $n\ell_z$ increases.

Finally, use dominant firms' R&D technology (10) on the BGP, that is, $\mathfrak{z} = \alpha n(\ell_z - \phi)$, and rewrite it as $\mathfrak{z} = \alpha(n\ell_z - \phi n)$, where both terms in the parenthesis are increasing in s, and do not explicitly depend on α. This shows that as α increases, so does $\alpha(n\ell_z - \phi n)$, and hence, $\mathfrak{z}$ increases. The last part assumes that $\mathfrak{z}$ is positive in equilibrium. Indeed, as mentioned before, we are only concerned with the symmetric equilibrium in which both R&D and product creation are positive. □

Proof of Proposition 5 In the model, the cost of goods sold (COGS), that is the labor used directly toward production, is $\ell_x - \psi$. Thus, we have $\ell_x - \psi = mc \times x$, where mc represents the marginal cost of production. Hence,

$$\ell_x - \psi = mc \times x = \frac{mc}{p} \times px = \frac{e-1}{e} \times px = \left(1 - \frac{1}{e}\right) px, \tag{B.34}$$

where the second-to-last equality follows since the ratio of the marginal cost to price is the inverse markup. Thus,

$$\frac{\text{COGS}}{\text{Sales}} = \frac{\ell_x - \psi}{px} = 1 - \frac{1}{e}, \tag{B.35}$$

and the gross profit ratio equals $\frac{1}{e}$ (i.e., $1 - \text{COGS}/\text{Sales}$). Since the price elasticity of demand $e = \epsilon - (\epsilon - 1)s$ decreases in s, the gross profit ratio falls in e and increases in s. Notably, the gross profit ratios is only a function of s and ϵ (i.e., it does not directly depend on cost parameters ϕ, ψ, and α). Thus, we can find the effect of cost parameters on the gross profit ratio using Proposition 3. In particular, since a rise in either ϕ or ψ increases s, then it also increases the gross profit ratio. In contrast, since a decline in α decreases s, it decreases the gross profit ratio. □

Proof of Proposition 6 The R&D-to-Sales ratio in the model is $\frac{\ell_z}{px}$, i.e., the fixed cost of R&D labs, ϕ, is part of the R&D expenditure. In particular, on the BGP, the expression for ℓ_z is in (B.32), and for the sales of dominant firms we have $px = \frac{s\mathcal{Y}}{N}$. Thus, dividing both sides of (B.32) by $\frac{s\mathcal{Y}}{N}$, we obtain

$$\frac{\text{R\&D}}{\text{Sales}} = \frac{\ell_z}{(\frac{s\mathcal{Y}}{N})} = \left[1 - (\rho + \delta)\gamma\right] \frac{1}{e} - \frac{\psi}{(\frac{s\mathcal{Y}}{N})}. \tag{B.36}$$

Notably, the right-hand side of (B.36) does not explicitly depend on ϕ or α but it depends on ρ, δ, γ, and s, and ϵ through e, and ψ from the last term, together with ω and θ through N. This enables us to show Parts (1) and (2) of the proposition through the effect of ϕ and α on s, as the following.

Part (1): In (B.36), ϕ affects the R&D-to-Sales ratio only through its effect on s. Since we know from Proposition 3 that ϕ increases s, then ϕ increases the R&D-to-Sales ratio.

Part (2): In (B.36), α affects the R&D-to-Sales ratio only through its effect on s. Since we know from Proposition 3 that a decline in α decreases s, then it must also decreases the R&D-to-Sale ratio.

Part (3): Since in (B.36), ψ directly affects the R&D-to-Sales ratio, discussing the effect of ψ using (B.36) is not easy. Instead, we use another expression for the R&D-to-Sales ratio that explicitly depends on ϕ and α but not ψ.

Note that from dominant firms' law of motion of technology (10) in the symmetric equilibrium, we have $\mathfrak{z} = \alpha n(\ell_z - \phi)$. Thus, $\ell_z = \frac{\mathfrak{z}}{\alpha n} + \phi$. Replacing $\mathfrak{z}$ from the dominant firms' investment rule (13) in the symmetric equilibrium, i.e., $\mathfrak{z} = -(\rho+\delta) + \alpha\theta\frac{\mathcal{Y}}{L}s\frac{e-1}{e}$, and then dividing both sides by $px = \frac{s\mathcal{Y}}{N}$, yields

$$\frac{\text{R\&D}}{\text{Sales}} = \Big(-\frac{\rho+\delta}{\alpha} + \phi n\Big)\frac{L}{s\mathcal{Y}} + \theta\frac{e-1}{e}. \tag{B.37}$$

Note that all three functions of s on the right-hand side, i.e., n, $\frac{1}{s\mathcal{Y}}$, and $\frac{e-1}{e}$ are decreasing in s. However, whether the product of the parenthesis and $\frac{1}{s\mathcal{Y}}$ is decreasing depends on the parenthesis being positive or negative. For instance, if $\phi = 0$, then the right-hand side of (B.37), written only in terms of s, simplifies to

$$-\left(\frac{\rho+\delta}{\alpha}\right)\frac{\epsilon-(\epsilon+\rho\gamma-1)s}{s\big[\epsilon-(\epsilon-1)s\big]} + \theta\frac{(\epsilon-1)(1-s)}{\epsilon-(\epsilon-1)s}. \tag{B.38}$$

It is not complicated but cumbersome to take the derivative of the above expression with respect to s, set it equal to zero, and solve for its extremums to find that the above sum has one maximum in $(0,1)$, whose expression is

$$s_\phi = \left(\frac{(\delta+\rho)(\epsilon-1) - \sqrt{(\epsilon-1)(\delta+\rho)\big[\alpha\theta-\gamma\rho(\delta+\rho)\big]}}{(\delta+\rho)(\epsilon+\gamma\rho-1)-\alpha\theta}\right)\Big(\frac{\epsilon}{\epsilon-1}\Big). \tag{B.39}$$

Thus, if ψ is sufficiently large that in equilibrium $s > s_\phi$, then the R&D-to-Sales ratio decreases in ψ; otherwise, the effect is the opposite.

On the other hand, if the parenthesis on the right-hand side of ($B.37$) is positive, which can happen if ϕ is large enough, then the right-hand side is decreasing in s since all the functions of s on the right-hand side are decreasing in s.

Finally, we want to mention that in our calibration, if $\phi = 0$, then in the range of s implied by the data, the R&D-to-Sales ratio is decreasing in almost two-thirds of the range, which is counterfactual. $\square$

Proof of Proposition 7 The proof has two steps. First, we show that s does not depend on L on the BGP. Second, we show that $\mathfrak{z}$ is a function of s (and other parameters) but does not depend on L.

First, to see that s on the BGP does not depend on L, start with ($B.27$), which determines s on the BGP. Note that all functions of s in ($B.27$) are: e, $\frac{L}{\mathcal{Y}}$, and $\frac{N}{\mathcal{Y}}$. Recall that $e = \epsilon - (\epsilon - 1)s$, which does not explicitly depend on L. Next, using (20) on the BGP, $\mathcal{Y} = \frac{L}{1-\rho\gamma\frac{s}{e}}$, and thus, $\frac{L}{\mathcal{Y}} = 1 - \rho\gamma\frac{s}{e}$, which again does not explicitly depend on L. Finally, on the BGP, $\frac{N}{\mathcal{Y}} = n(1 - \rho\gamma\frac{s}{e})$, where $n = \frac{N}{L}$. However, that from (21), n depends on s but not L. As a result, ($B.27$) is an implicit function of s, which does not depend on L, i.e., s on the BGP does depend on L.

Second, to observe that $\mathfrak{z}$ on the BGP is a function of s but not L, see ($B.30$), which is an exact expression $\mathfrak{z}(s)$ on the BGP. $\square$

C. Equilibrium Conditions and Targeted Moments

The following system of differential equations determines the model's equilibrium:

$$e = \epsilon - (\epsilon - 1)s; \tag{C.1}$$

$$n = \omega^{\frac{-\epsilon}{(\epsilon-1)\theta}} \left(\frac{e-1}{e}\right)^{\frac{1}{\theta}} \left(\frac{1-s}{s}\right)^{\frac{1}{(\epsilon-1)\theta}}; \tag{C.2}$$

$$\mathcal{Y} = \frac{L}{1 - \gamma\frac{s}{e}(\rho + \frac{\dot{e}}{e} - \frac{\dot{s}}{s})}; \tag{C.3}$$

$$r = \frac{\dot{\mathcal{Y}}}{\mathcal{Y}} + \rho; \tag{C.4}$$

$$\mathfrak{z} = -(r+\delta) - \frac{\dot{n}}{n} + \alpha\theta\frac{\mathcal{Y}}{L}s\frac{e-1}{e}; \tag{C.5}$$

$$\frac{\dot{n}}{n}\Big(1 - \frac{eL}{\alpha\gamma s\mathcal{Y}}\Big) + \frac{\dot{e}}{e} - \frac{\dot{s}}{s} = \frac{1-\theta(e-1)}{\gamma} + \Big(r+\delta\Big)\frac{eL}{\alpha\gamma s\mathcal{Y}} - \frac{enL}{\gamma s\mathcal{Y}}(\psi + \phi). \tag{C.6}$$

Equations (C.1)-(C.6) correspond to (11), (21), (20), (19), (22), and (23).

The targeted moments are

$$\text{gross profit ratio} = \frac{1}{e}, \tag{C.7}$$

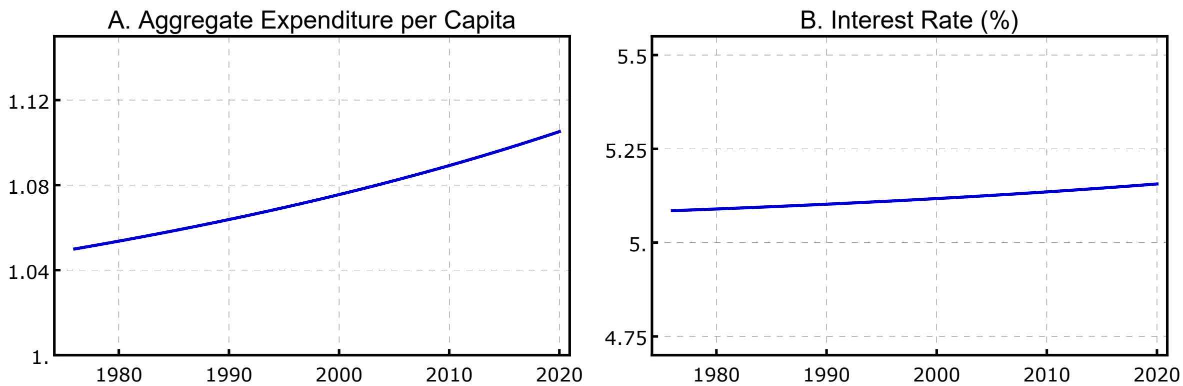
$$\frac{\text{R\&D}}{\text{Sales}} = \frac{1}{e}\left[1 - \gamma(\delta+\rho) + \Big(\frac{\dot{s}}{s} - \frac{\dot{e}}{e} - \frac{\dot{n}}{n}\Big)\gamma\right] - \frac{\psi n L}{s\mathcal{Y}}, \tag{C.8}$$

$$\text{TFP growth rate} = \mathfrak{z} + \big(1 - \theta s\big)\frac{\dot{n}}{n}. \tag{C.9}$$

To proceed, we restrict the time path of the market share of dominant firms, $s(t)$, to be a quadratic time trend, making optimization manageable. Given the differential equation for $s(t)$ in (C.6), this approach is equivalent to assuming a time path for the exogenous driving force ϕ we aim to back out from the data. Note also that we do not need to make extra assumptions for the initial values $s(0)$ and $\dot{s}(0)$ as the assumed quadratic path already pins them down. Assuming an explicit functional form for the path of $s(t)$ allows us to write explicit functions for the first and second-order time derivatives. This, in turn, allows us to write transition paths for our target dynamic moments, i.e., gross profit ratio, R&D-to-sales ratio, and TFP growth. Thus, each set of parameters φ_0 and φ_1, along with the assumption of a quadratic time of path of $s(t)$, and the model parameters, gives us a unique time path for our target moments, which then we fit to the data. Minimizing a weighted sum of squared errors gives the calibrated parameter values.

D. Additional Results

Figure *D*.1: Interest Rates, 1976-2020



Notes: The figure shows model simulated paths with perfect foresight for aggregate expenditure per capita (panel A) and the interest rate (panel B).

Figure *D*.1 shows that aggregate expenditure per capita (panel A) grows at a roughly constant rate throughout the period 1976-2020, implying a virtually constant interest rate (panel B) that is broadly consistent with the after-tax rate of return to private businesses in the United States (Gomme, Ravikumar and Rupert, 2011).