# Turbulent Growth:

# Business Dynamism and Aggregate Productivity*

(Job Market Paper: Click here for the latest version.)

Filippo Massari†

Department of Economics

North Carolina State University

November 6, 2021

**Abstract**

This paper formulates a model of turbulent endogenous growth. Turbulence denotes the endogenous job reallocation due to entry, exit, and churning (movements within the firm-size distribution). The insight central to the model is that the forces that drive aggregate growth also drive turbulence because the two are manifestations of a single underlying process: competition for market share.

The profit motive drives competition for market share through R&D. When firms expand their market share they face a lower relative price, reducing the marginal value of further gains in market share. This leads to the emergence of diminishing returns in relative terms. Therefore, incentives to innovate decline in relative size, generating churning endogenously as mean-reversion. This mechanism delivers a stationary, non-degenerate firm-size distribution with a realistic right tail dependent on R&D. Meanwhile, constant returns to the cumulative factor (knowledge) drive a sustained aggregate growth rate determined by R&D. Endogenous entry and exit entail selection effects that shape the characteristics of the firm population, and generate a firm life cycle, affecting R&D, thus growth.

In a quantitative application that replicates changes in business dynamism in the US, I find that aggregate productivity growth declines mildly despite higher R&D effort. (*JEL*: E19, L11, O14, O31, O41)

***Keywords***— Aggregate Productivity, Firm Dynamics, Turbulence, Endogenous Growth, Firm-Size Distribution, Heterogeneous Agents, Business Dynamism, Market Share.

*Acknowledgement: I am grateful to my advisors Pietro Peretto and Giuseppe Fiori for their constant guidance and encouragement. I have also benefited greatly from various discussions with Domenico Ferraro, James Nason, Daisoon Kim, Soroush Ghazi, and Ayse Kabukcuoglu Dur. I thank all parcipants to the Macroeconomics Lunch Seminars at North Carolina State University. Lastly, I acknowledge the technical support of the High-Performance Computing Services at North Carolina State University. All remaining errors are my own.

†Nelson Hall 4129, Box 8110, NCSU Campus, Raleigh, NC 27695. E-mail: fmassar@ncsu.edu. Website: http://filippomassari.com/.

# 1 Introduction

Aggregate productivity growth is a *turbulent process*, intrinsically connected to business dynamism. It is the key to improving living standards and among the essential factors that economists strive to understand. Modern growth theory and empirics have emphasized R&D as the driver of productivity growth. Importantly, they point out that firm-level R&D is responsible for firm growth and firm formation, while lack of it may ultimately lead to firm death.

What is a *turbulent process*? The business environment is characterized by high entry and exit rates. Additionally, firms expand and contract as they gain or lose market share, a phenomenon known as churning, which entails reshuffling of firms' position within the firm-size distribution. Brown et al. (2008, p. 3) define turbulence to include both these aspects and identify an appropriate measure for it, namely, the job reallocation rate. They write that turbulence is "the entire process of economic change: worker reallocation as workers change jobs and job reallocation from firms contracting and shutting down, to firms expanding and starting up."

Discussions of turbulence and business dynamism are currently prominent due to their decline in the past few decades in the US.[1] Meanwhile, despite the increase in aggregate R&D, the growth rate of aggregate productivity has failed to rise (Byrne et al., 2016; Syverson, 2017; Fernald, 2018). Nevertheless, the link between all the dimensions of business dynamism and aggregate productivity growth is hard to rationalize.

The main goal of this paper is to first study the firm formation, firm growth, and firm death processes that drive key aspects of business dynamism, and then link them to aggregate productivity growth. Specifically, differences in growth rates across firms display systematic elements, such as negative correlation with initial relative size (Sutton, 1997; Caves, 1998; Audretsch et al., 2004), higher frequency of exit among small and unproductive firms, and negative correlation between firms' growth rates and age (Haltiwanger et al., 2012). Additionally, I address two questions through an application of the model. First, what consequences do changes that explain declining business dynamism have for aggregate labor productivity growth? Second, do these changes reproduce the productivity growth slowdown observed in the past couple of decades relative to the

[1] The main trends observed in the data include a decrease in the job reallocation rate since 2000 (Decker et al., 2020), a reduction in the startup rate (Decker et al., 2014), especially of high growth firms (Decker et al., 2016), and, as a consequence, an increase in the average firm age Hopenhayn et al. (2018).

previous one? The debate over this question is ongoing and still open.[2]

To achieve these goals, I build on the endogenous growth and the firm dynamics literatures. I develop a framework with a continuum of firms,[3] which produce unique but substitutable goods (monopolistic competition). They perform innovation in-house to enhance their stock of technological knowledge through which they reduce their marginal cost of production, hence their price. By reducing their price, firms steal market share from their competitors. Entry and exit decisions are endogenous. In this framework, the fundamental source of heterogeneity is an idiosyncratic stochastic R&D productivity (ability to innovate) drawn from a common distribution. Aggregate variables evolve deterministically.

This paper contributes to the literature in four main ways. The first contribution is to show that churning is at least partly endogenous and a fundamental property of the process of aggregate productivity growth. In particular, the model identifies a mechanism through which R&D differs across firms generating differences in growth rates, causing churning to arise. The possibility to steal market share from others provides incentives to innovate. Simultaneously, this competitive force gives rise to *economic diminishing returns*, as increasing one's own market share reduces the remaining market share that can be gained with further innovation. Consequently, incentives to innovate decline with market share giving rise to negative scale dependence — the negative correlation between firms' growth rates and their relative size — in the form of mean-reversion. However, as these diminishing returns are in *relative* terms, they do not disrupt sustained growth, which requires constant or increasing returns to the growth-driving factor. As a result, the aggregate productivity growth rate depends on firms' optimal R&D decisions as in other fully endogenous growth models.

The theoretical ingredients responsible for this scale dependence are standard: (i) in-house R&D; (ii) an idiosyncratic stochastic R&D productivity that is the root cause of heterogeneity in firms' knowledge stock and market share; and (iii) imperfect substitutability between goods, which introduces demand-driven diminishing returns to relative knowledge — the firm's knowledge stock relative to the average knowledge stock. As firms face a downward-sloping demand curve, expanding their volume of production relative to their competitors requires reducing their

[2]For a thorough summary see Naude (2020).

[3]Although this paper's focus is on the product-line as this is the relevant unit for discussing aggregate productivity growth, I use the term firm as a synonymous throughout the paper.

relative price. This exerts downward pressure on the return to further market share expansion, thus reducing the incentive to innovate and grow even larger.

A second related contribution concerns the firm-size distribution. Through the economic mechanism illustrated above, I point to a novel source of stationarity, and I characterize its right tail. As firms' return to R&D investment, all else constant, declines in their level of relative knowledge, this model gives rise to mean-reversion in relative knowledge (hence in size). In this way, it introduces a force that prevents firms from perpetually growing bigger or shrinking. Furthermore, as documented by Rossi-Hansberg and Wright (2007), the right tail of US firms and establishments declines faster than the one of a Pareto or lognormal distribution. My model produces a firm-size distribution arising from firms' dynamic optimization that is non-degenerate and consistent with this fact. As firms grow large, their R&D effort will tend towards zero, thus eliminating any possibility of growing above a certain size. This mechanism creates an upper bound for firm sizes that is manifested in a decline of the distribution's right tail. Therefore, the shape of the distribution depends on the parameters that regulate firms' incentives to innovate and the speed of mean reversion. That is, any parameter change that affects firms' investment decisions would modify the distribution's right tail, turbulence, and the economy's growth rate.

As a third contribution, I offer a take on the consequences of changing business dynamism through a quantitative counterfactual analysis based on the model. The model calibrated on US data suggests that the causes typically associated with changing business dynamism do not explain the productivity growth decline observed comparing the periods '94-'04 to '05'-'15. Instead, in line with the data, these changes generate an increase in R&D without a proportional increase in the aggregate productivity growth rate, which instead is mildly decreasing.

This result arises because an increase in factors that prevent entry (in my experiment, an increase in the sunk entry cost) have two major effects. First, they increase the average market share. Therefore, R&D rises because of a stronger cost spreading effect — namely the ability to spread the cost of R&D on more units sold since knowledge is non-rival. Second, they shield poorly performing firms from the competition of entrants that would force them to exit. As these firms are small, their return to R&D is higher than their larger counterparts since they face a higher relative price. But, as these firms are not very innovative, they expand their knowledge frontier by little, thus having a negligible impact on aggregate productivity. Indeed, the reason why the growth rate of aggregate productivity does not rise is that the composition of firms in the market is different. Specifically, there are fewer entrants and young firms — that on average have a higher ability to innovate. Overall, even though there is more R&D on aggregate, firms that perform R&D tend to be less innovative.

The final significant contribution is the framework, which allows studying entry, exit, churning, the number of goods, the firm-size distribution, and aggregate productivity growth jointly. This paper merges the literature on firm dynamics with the endogenous growth literature, specifically Hopenhayn (1992) and Peretto and Connolly (2007).

As in Hopenhayn (1992), the model delivers endogenously entry, exit, and firm dynamics. These dynamics are driven by idiosyncratic shocks to which firms respond actively by adjusting their size. Contrary to Hopenhayn, where the shock hits firms' productivity, in my model, differences in the productivity level arise endogenously as the outcome of firms' R&D investment. This deviation allows me to derive an endogenous aggregate productivity growth rate jointly determined with the other moments of the model. The framework proposed by Hopenhayn is the foundation of the firm dynamics literature reviewed in Hopenhayn (2014), which has recently devoted much attention to resource allocation and aggregate productivity. By introducing endogenous aggregate productivity growth, I contribute to advancing the research agenda put forth by Restuccia and Rogerson (2017, p. 168). In their words: "[f]rom a modeling point of view, the key issue is to extend the simple static model of heterogeneous producers [...] to a dynamic setting that includes endogenous decisions that influence future productivity", to "go beyond static effects of misallocation, and focus on the potentially much larger dynamic effects."

As in Peretto and Connolly (2007), the theory includes vertical innovation — cost reduction in the production of existing goods — and horizontal innovation — development of new products — where the former is the engine of long-run growth while the latter drives the equilibrium number of goods. While that model focuses on the symmetric equilibrium, my model introduces an idiosyncratic shock to firms' R&D outcome, thus delivering a non-degenerate firm-size distribution and richer firm dynamics by giving rise to churning and allowing for simultaneous entry and exit.

Endogenous growth models with heterogeneous producers are not new. The intellectual foundations were laid down in the field of industrial organization, specifically Ericson and Pakes (1995). Notable examples of general equilibrium analyses that give rise to endogenous aggregate productivity growth date back to Thompson (2001) and Laincz (2009).[4] Contrary to my work, the former paper assumes away the economic diminishing returns to endogenous productivity, thus removing any dependence of R&D on firm-specific market share. In my model, this is an important driver of turbulence, and an element that adds a feedback mechanism from the firm-size distribution to R&D investment, therefore to aggregate productivity growth. Furthermore, my paper introduces exit as an optimal stopping problem. Laincz (2009) delivers a tendency towards monopoly countered only by technological diffusion from the industry leader to entrants. My model, instead, obtains a non-degenerate distribution from assuming product differentiation, and it is therefore complementary to Laincz's model as the two frameworks describe different types of market.

Finally, delivering a stationary and non-degenerate firm-size distribution has so far been a tough theoretical challenge. Most models lack a force of attraction that can reproduce it. They either require creative destruction from entrants to prevent firms from becoming too small or an exogenous death shock that prevents the most successful firm from becoming too large and monopolizing the market (Klette and Kortum, 2004; Acemoglu and Cao, 2015; Cao et al., 2017; Acemoglu et al., 2018; Akcigit and Ates, 2019). My model, instead, features an internal mechanism that generates this force of attraction, while the inclusion of entry and exit is unnecessary for this purpose. Other models use knowledge spillovers that facilitate the growth of smaller firms (Thompson, 2001; Laincz, 2009). Without dismissing the importance of entry, exit, and knowledge spillovers, I point to an additional force that delivers a stationary, non-degenerate, and realistic firm-size distribution.

## 2 A Model of Turbulent Growth

This section describes the model. Time is discrete. A monopolistically competitive intermediate sector consists of a mass of firms that produce a unique good, sold to a perfectly competitive fi-

[4] The relevance of endogenous growth models with heterogeneous producers is corroborated by the proliferation of studies, primarily based on Klette and Kortum (2004), that focus on understanding a variety of phenomena or informing policy (Acemoglu et al., 2018; Mukoyama and Osotimehin, 2019; Akcigit and Kerr, 2018). These papers motivate heterogeneity by pointing out that entrants and mature firms have different abilities to innovate. As a result, heterogeneity plays a role through a selection effect only: parameter changes that affect the entry and exit rates modify the average ability to innovate in the economy, thus R&D incentives, and growth. My model includes this aspect, but it goes a step further as different firm-size distributions imply different incentives to innovate.

nal sector that assembles them in a final good. The final sector distributes the final good to the representative household for consumption, and to entrepreneurs for setting up new firms. Innovation takes up two different forms: the *technological depth* is augmented by improving production processes for existing goods; the *technological breadth* is expanded through the introduction of new goods. Firms invest in R&D to lower their goods' production cost and face an endogenous exit decision at the end of the period. The presence of firm-level idiosyncratic uncertainty over R&D's outcome generates heterogeneity in firms' productivity levels. New firms enter the market upon payment of a sunk cost in units of output by introducing a new good. All the aggregate variables evolve deterministically.

Households face a consumption/saving choice, along the lines of Bilbiie et al. (2012). Furthermore, they supply labor inelastically.

## 2.1 Households

The economy is populated by a representative household of size $L_t = L_0(1+\lambda)^t$, where $\lambda$ is the population growth rate. The household is endowed with $L_t$ units of labor that it supplies inelastically. It makes decisions on how to allocate its income to consumption goods or saving at each point in time.

The representative household maximizes its lifetime utility function,

$$\max_{\{c_t,s_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t L_t \frac{c_t^{1-\varphi}-1}{1-\varphi}, \tag{1}$$

by choosing the sequence of per capita consumption in the final good, $c_t$, and their saving in a portfolio of stocks of real value $s_{t+1}$. $\varphi > 0$ is the inverse of the intertemporal elasticity of substitution.

The household derives its income from the per capita real wage $w_t$, and a return $r_t$ on the portfolio of stocks, while it allocates this income to consumption and saving in the portfolio itself. As in Bilbiie et al. (2012), the portfolio is managed by a risk-neutral manager who operates in a perfectly competitive environment. It includes all firms that populate the economy and new firms, whose entry cost is financed by issuing equity. This implies that the idiosyncratic risk is diversified away, simplifying the problem. The household faces the following budget constraint expressed in real terms:

$$s_t + c_t L_t \leq (1 + r_t)s_{t-1} + w_t L_t. \tag{2}$$

Combining the first-order conditions, I obtain the Euler equation that governs the household's saving decision,

$$1 + r_{t+1} = \frac{1}{\beta}\left(\frac{c_t}{c_{t+1}}\right)^{-\varphi}. \tag{3}$$

## 2.2 Final Sector

I now turn to the description of the economy's production side, starting from the final sector to derive the demand for each intermediate good. A perfectly competitive final sector sells the final good to the household and to entrepreneurs who need it to finance the sunk entry cost. It assembles the final good according to a CES aggregator:

$$Y_t = \left[\int_0^{N_t} x_{i,t}^{\frac{\epsilon-1}{\epsilon}} di\right]^{\frac{\epsilon}{\epsilon-1}}, \tag{4}$$

given a real output $Y_t$, made from units of the different intermediate goods $x_{i,t}$, the only inputs. $N_t$ is the mass of goods, and $\epsilon > 1$ is the elasticity of substitution across them. The price index is:

$$P_t = \left[\int_0^{N_t} P_{i,t}^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}}, \tag{5}$$

where $P_{i,t}$ is the price of each good $i$.

The representative retailer maximizes his profits by supplying the household and potential entrants with units of the basket of goods. The profit maximization yields the following demand schedule for good $i$:

$$x_{i,t} = Y_t p_{i,t}^{-\epsilon}, \tag{6}$$

where $p_{i,t} = \frac{P_{i,t}}{P_t}$ is the relative price.

### 2.3 Intermediate Sector: Production, Innovation, Entry, and Exit

This subsection describes the intermediate sector of the economy. It consists of incumbents, entrants, and exiting firms.

At time $t$, the intermediate sector is populated by $N_t$ firms with market power producing their own unique good.

The demand schedule derived above implies a revenue per good of:

$$\underbrace{P_{i,t}x_{i,t}}_{\text{Revenue}} = \underbrace{P_t Y_t}_{\substack{\text{Market}\\\text{size}}} \underbrace{p_{i,t}^{1-\epsilon}}_{\substack{\text{Market}\\\text{share}}}, \tag{7}$$

which can be decomposed into market size and market share.[5] The decomposition provides an insight into the competitive process underlying the model. As $\epsilon > 1$, firms can gain market share at others' expense by lowering their relative price. Additionally, two opposite forces affect revenue per good: changes in aggregate spending, namely market size, and changes in the number of producers, which dilute market shares. Market size is beyond the control of the firm, therefore the only way to increase their revenue is for the firm to reduce price and steal market share from others.

The following subsections describe, in turn, the decisions of incumbents and entrants.

#### 2.3.1 Incumbents

Incumbents face a demand given by equation (6). They employ labor that is allocated to produce the intermediate good, $l_{x_{i,t}}$, to cover the fixed costs of production $\Phi$, and to produce knowledge that reduces the future cost of production, namely to perform R&D, $l_{Z_{i,t}}$. They maximize their value, which is the present value of the stream of dividends, by choosing the optimal price, production labor, R&D labor, and whether to exit the market or not.

For the sake of exposition, I break their optimization problem into a static and a dynamic component to derive a cleaner Bellman equation as in other related works, such as Acemoglu et al. (2018). The static component is a per-period dividend maximization, holding constant R&D

[5] By rearranging equation (7) to isolate $p_{i,t}^{1-\epsilon}$, one can observe that it equals the ratio of expenditure on good $i$ and total expenditure, the definition of market share.

investment. This allows me to derive an optimal operating profit, conditional on the state, that can be plugged into the Bellman equation. The dynamic component involves an investment decision to maximize the firm's value, with an option to exit the market if it turns negative.

This structure follows Peretto and Connolly (2007), except for returns to R&D that are uncertain. When the investment decision is taken, the firm does not know precisely how much new knowledge will be generated by the research effort. This idiosyncratic uncertainty generates heterogeneity in the knowledge stock across firms, and, consequently, in their productivity level.

#### Timing of Events

The timing of the events is the following: first, an incumbent firm observes its draw of the parameter that determines the outcome of its past R&D investment, thus the realization of $Z_{i,t}$, its knowledge stock. Second, it hires labor to produce, invest in R&D, and cover its fixed cost; it picks the price and sells its good. After that, it distributes dividends to the household.[6] Finally, at the end of the period, it decides to exit if its continuation value is negative.

#### Static Problem: Profits

In each period, dividends are given by

$$\pi_{i,t} = p_{i,t}x_{i,t} - w_t(l_{x_{i,t}} + l_{Z_{i,t}} + \Phi). \tag{8}$$

where $l_{x_{i,t}}$ is labor allocated to production, and $\Phi$ is overhead labor. As mentioned above, to simplify the exposition, I break down the firm maximization problem into a static and a dynamic component. This is equivalent to firms maximizing dividends in each period by choosing how much to produce and what price to charge, holding for the moment R&D labor, $l_{Z_{i,t}}$, constant.

Following the literature, the production technology includes only productivity and labor, such that:

$$x_{i,t} = Z_{i,t}^{\theta} l_{x_{i,t}}^{\nu} \qquad 0 < \theta, \nu < 1, \tag{9}$$

[6] As there are no liquidity constraints, a negative dividend would imply borrowing from the household through the financial intermediary.

where $Z_{i,t}$ is the endogenous stock of knowledge possessed by firm $i$, and the parameter $\theta$ determines the returns to knowledge, or the extent to which production is knowledge-intensive.

The static maximization problem requires a choice of production labor, a price and a quantity to maximize equation (8), subject to demand (6), and the production function (9). The first order conditions yield a production labor demand of:

$$l_{x_{i,t}} = \left[ \nu \frac{\epsilon - 1}{\epsilon w_t} \left( \frac{y_t L_t}{N_t} \right)^{\frac{1}{\epsilon}} Z_{i,t}^{\theta \frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon - \nu(\epsilon-1)}}. \tag{10}$$

Firms' production labor demand is increasing in the productivity level, $Z_t^{\theta}$, and decreasing in wage. It increases with the overall spending on final goods and declines in the number of goods. In other words, it increases in market share.

A labor demand schedule as in (10) implies a price of:

$$p_{i,t} = \frac{\epsilon}{\epsilon - 1} \frac{w_t l_{x_{i,t}}^{1-\nu}}{\nu Z_{i,t}^{\theta}}. \tag{11}$$

This optimal pricing strategy involves charging a constant markup over marginal cost. Moreover, the marginal cost is a function of production labor, as diminishing returns to production labor are present. Importantly, firms can reduce their relative price by improving their technological knowledge.

As anticipated earlier, from equation (7), reducing the relative price, thus gaining market share, is the only way for firms to increase their revenue. Therefore, equation (11) illustrates the fundamental way in which competition occurs: by accumulating technological knowledge faster than the rate of wage growth, firms can lower their relative price and steal market share from competitors. In other words, firms have an incentive to innovate because they can gain market share at the expense of others, and increase their revenue as a result.

Substituting (10) and (11) into equation (8) and using equation (9), dividends can be re-expressed as a function of $Z_{i,t}$, and $l_{Z_{i,t}}$ only.

#### Heterogeneity and Dynamics: Firm Value Maximization and Exit Decision

Here, I present the dynamic problem of the firm. Each firm makes an investment decision to increase their future knowledge, thus reducing their production cost. The outcome of the R&D investment is subject to idiosyncratic uncertainty, driving heterogeneity in firms' productivity level.

Firms increase their future stock of knowledge through R&D investment. Following Peretto and Smulders (2002), the R&D technology is:

$$Z_{i,t} - Z_{i,t-1} = \alpha_{i,t} Z_{i,t-1}^{\mu} K_{t-1}^{1-\mu} l_{Z_{i,t-1}}^{\zeta}, \tag{12}$$

$0 < \mu < 1$ is a parameter that regulates the private and social returns to knowledge, $\alpha_{i,t} > 0$ is the firm-specific productivity of R&D, and $K_t$ is the knowledge spillover, namely, the element that captures the partial non-excludability of knowledge, and the consequent ability of firms to make use of knowledge acquired by others. It takes up the following form:

$$K_t = \frac{1}{N_t} \int_0^{N_t} Z_{i,t} di \equiv Z_t. \tag{13}$$

An R&D technology of this kind captures four important elements. First, new knowledge is a function of the existing stock of knowledge due to its cumulative nature, i.e. new knowledge builds on existing knowledge. The linearity is the simplest and most tractable specification in which knowledge is the factor that drives long-run exponential growth at a constant rate.[7]

Second, only the firm that produces good $i$ possesses the expertise to improve that line of production, based on the idea that a large driver of innovation is firm-specific in-house technology, widely documented empirically (Dosi, 1988; Garcia-Macia et al., 2019). Therefore, R&D is performed in-house, implying that firms know what product-line they are improving upon when the investment decision is taken. The presence of this element is necessary to deliver the scale dependence in growth rates that generates the mean reversion that produces churning and makes the firm-size distribution stationary.

Third, in line with empirical evidence (Laincz and Peretto, 2006; Ha and Howitt, 2007; Madsen, 2008), the spillover occurs from average knowledge and does not increase with the number of different goods produced in the economy. This specification incorporates the idea that the technological distance between lines of research increases as the product market grows larger, thus

[7] Peretto (2018) provides a generalization that allows for new knowledge to exhibit increasing returns in the existing stock of knowledge.

diluting away the knowledge spillover and eliminating the scale effect, as shown in previous works (Peretto, 1998; Dinopoulos and Thompson, 1998; Young, 1998).

Fourth, the firm-specific shock $\alpha_{i,t}$ follows an AR(1) process:

$$\begin{aligned} \log \alpha_{i,t} = (1-\rho)\log \overline{\alpha} + \rho \log \alpha_{i,t-1} + \xi_{i,t}, \qquad \xi_{i,t} \sim N(0, \sigma_\xi) \\ 0 \leq \rho < 1 \end{aligned} \tag{14}$$

where $\xi_{i,t}$ is the draw and $\sigma_\xi$ its standard deviation. The idiosyncratic productivity shock is the only unknown variable to the firm manager when the investment decision is taken, in line with empirical evidence (Doraszelski and Jaumandreu, 2013). This is the fundamental source of heterogeneity in this model.

At the beginning of each period, after observing the draw, each firm invests to maximize its value:

$$\max_{\{l_{Z_{i,t+h}}, Z_{i,t+1+h}\}_{h=0}^{\infty}} V_{i,t} = \pi_{i,t}(l_{Z_{i,t}}, Z_{i,t}) + \max\left\{\mathbb{E}_t \sum_{h=1}^{\infty}\prod_{q=1}^{h}\frac{1}{1+r_{t+q}}\pi_{i,t+h}(l_{Z_{i,t+1}}, Z_{i,t+1}), 0\right\} \tag{15}$$

Future profits are discounted using the risk-free interest rate $r$, which is determined by the representative household's time preferences.

The dynamic optimization can be re-expressed as a Bellman equation:

$$V(Z_{i,t}, \alpha_{i,t}) = \max_{\{l_{Z_{i,t}}\}}\left\{\pi_{i,t}(Z_{i,t}, l_{Z_{i,t}}) + \frac{1}{1+r_{t+1}}\max\{\mathbb{E}_t V(Z_{i,t+1}, \alpha_{i,t+1}), 0\}\right\} \tag{16}$$

constrained by the knowledge accumulation equation (12).

#### 2.3.2 Entry: Creation of New Goods

I now turn to the description of the entry decision. Entry occurs as long as the present value of the expected stream of dividends exceeds the sunk cost of setting up a firm.

Entrants issue equity to finance the cost of entry. The payment of the sunk cost is in units of output.

When taking the entry decision, entrepreneurs know that their knowledge level in the following period will be drawn out of a lognormal distribution (as the firm size distribution is skewed in the data) around the average knowledge level in the economy. Entry knowledge is given by:

$$Z_{t+1}^d \sim Lognormal(\chi_Z, \sigma_Z^E) Z_{t+1}. \tag{17}$$

Entrepreneurs will also draw their initial R&D productivity from:

$$\alpha_{t+1}^d \sim Lognormal(\chi_\alpha, \sigma_\alpha^E). \tag{18}$$

The distribution is lognormal as entrants display skewness in their employment growth rates (Decker et al., 2016).

As all potential entrants draw from the same distributions, their value is the same. Given an expected initial level of productivity and productivity of R&D, there is entry at time $t$ as long as:

$$v_t^E = \mathbb{E}_t V(\alpha_{t+1}^d, Z_{t+1}^d) \geq Z_t^\theta N_t^{\frac{1}{\epsilon-1}} f_E, \tag{19}$$

where the right side of the inequality is the entry cost made up of a fixed component $f_E$ and of the technological depth, $Z_t^\theta$, and breadth, $N_t^{\frac{1}{\epsilon-1}}$, of the economy. This specification has a practical purpose: in a growing economy, the entry cost must scale with everything else. Otherwise, as the economy grows richer, setting up new firms would become cheaper, introducing a trend in the entry rate which would be counterfactual. The specification presented here is the simplest one consistent with this property, but not the only one that can deliver it. The idea captured by this specification is that with an increase in the sophistication of the production techniques and of the variety of goods available, the capital required to set up a firm increases proportionally.

Firms set up at time $t$ face the same problem as incumbents in period $t + 1$.

## 2.4 Equilibrium

The equilibrium of the model is defined by:

- a wage $w_t$, interest rate $r_t$ and price index (5) that firms and the household take as given;
- a demand function (6) from the final sector for the intermediate goods;
- a labor supply $L_t$, and a demand function for production, overhead and R&D labor;

- an Euler equation (3) for the representative household;
- the free entry condition (19);
- a law of motion of firms:

$$N_{t+1} = N_t + N_{E_t} - N_{X_t}; \tag{20}$$

- a value function $V(Z_{i,t})$;
- and a distribution $\Gamma_t(z_t)$ of relative knowledge, $z_{i,t}$, where $z_{i,t} = \frac{Z_{i,t}}{Z_t}$,

such that the following conditions hold.

First, the interest rate adjusts to guarantee that the value of the portfolio held by the household equals the aggregation of the value of all firms:

$$s_t = \int_0^{N_t - N_{X_t}} v_{i,t} di + \int_0^{N_{E_t}} v_{i,t}^E di. \tag{21}$$

Second, exploiting equation (5), the prices for each variety are such that they guarantee goods market-clearing:

$$Y_t = c_t L_t + Z_t^{\theta} N_t^{\frac{1}{\epsilon-1}} f_E N_{E_t}. \tag{22}$$

Third, the wage adjusts to ensure that quantity of labor demanded by each firm for each activity equals its inelastic supply:

$$L_t = \int_0^{N_t} (l_{x_{i,t}} + l_{Z_{i,t}}) di + \Phi N_t. \tag{23}$$

### 2.4.1 Steady State

I solve the model for the stationary steady state equilibrium numerically. I present the stationarized version in appendix A. Appendix B includes a description of the algorithm used to solve the model.

There exists a time-invariant distribution of firms over relative knowledge $\Gamma(z)$ in steady state that is unique given any initial distribution. The following section describes the forces that make this distribution unique and stationary.

Before introducing output growth, it is useful to define relative (to the arithmetic average) knowledge

$$z_{i,t} = \frac{Z_{i,t}}{Z_t}, \tag{24}$$

and the number of firms per capita

$$n_t = \frac{N_t}{L_t}, \tag{25}$$

which is also the inverse of average firm size and remains constant in steady state, where $N_t/N_{t-1} = 1 + \lambda$. For this class of models, the stationarity of average firm size is discussed in Peretto and Connolly (2007). The basic insight is that as population increases, the market size gets higher, thus increasing operating profits. Larger profits stimulate entry. Entry drags down the average market share, restoring the orginial profit level at the original average firm size.

Furthermore, to simplify the notation, define productivity as:

$$A_{i,t} = Z_{i,t}^{\theta}. \tag{26}$$

**Aggregate Productivity Level**

From the CES aggregator given by equation (4) and the production function in equation (9), I can express real output per capita as:

$$\frac{Y_t}{L_t} = \underbrace{\underbrace{N_t^{\frac{1}{\epsilon-1}}}_{\substack{\text{Tech.}\\ \text{breadth}}} \underbrace{A_t}_{\substack{\text{Tech.}\\ \text{depth}}} \underbrace{\left\{ \frac{1}{N_t} \int_0^{N_t} \left[ (S_{i,t} N_t)^{\nu} a_{i,t} \right]^{\frac{\epsilon-1}{\epsilon}} di \right\}^{\frac{\epsilon}{\epsilon-1}}}_{\substack{\text{Allocative}\\ \text{efficiency}}}}_{\text{Aggregate productivity}} \underbrace{\frac{L_{x_t}}{L_t} \left( \frac{L_{x_t}}{N_t} \right)^{\nu-1}}_{\text{Production effort}}, \tag{27}$$

where $S_{i,t}$ is the production labor share for firm $i$. Output per capita can be thought of as a combination of productivity and resources devoted to production. The latter element depends on the total fraction of labor devoted to producing units of the intermediate goods — where $L_{x_t}$ denotes aggregate production labor —, and on the average firm size due to the presence of diminishing returns to production labor. I focus on this model-consistent definition of aggregate productivity because, as the ultimate interest of any analysis on economic growth is the increase in household's utility, the relevant unit to consider is the output of the final good — which partly goes to consumption — and the effort exerted to produce it. Other definitions of productivity correlate with this one.

Importantly, all terms that make up the aggregate productivity level are endogenous and depend exclusively on a vector of relative knowledge levels $z_t$ and R&D productivity $\alpha_t$.

A contribution of this paper is to decompose the aggregate productivity level into various factors that can be linked to firm-level productivity. Due to non-linearities, the dispersion in productivity levels and firm sizes is manifested in the aggregate productivity level. This decomposition shows explicitly how. Aggregate productivity includes three different elements. $N_t^{\frac{1}{\epsilon-1}}$ is the love of variety effect implied by the CES aggregator. This arises out of product differentiation and a preference structure that rewards a larger variety of goods in the market.

The second term describes the technological depth of the economy, and it corresponds to the average productivity across firms, defined as:

$$A_t = \int_0^{N_t} S_{i,t} A_{i,t} di, \tag{28}$$

As I will clarify later, this definition of average productivity is useful to understand the role that heterogeneity plays in shaping the aggregate productivity level. Coincidentally, this definition is the one commonly adopted in empirical studies (Foster et al., 2001; Melitz and Polanec, 2015). While in those papers the choice is arbitrary, this model offers a theoretical justification for it, but at the same time it points out its limitations by making explicit the role of higher order moments of the relative productivity distribution.

Finally, the last term shows that aggregate productivity depends on the distribution of weighted firm relative productivities, as:

$$a_{i,t} = \frac{A_{i,t}}{A_t}. \tag{29}$$

This element describes the allocative efficiency of the economy. The distribution of individual productivities matters for aggregate productivity for two reasons, both having to do with aggregation. First, the CES aggregator is a power mean, which is altered by the firms' relative productivity distribution. Furthermore, the presence of diminishing returns in production interferes with the weights (here expressed as the labor share scaled by $N_t$). To understand why this term is tied to the distribution of productivities and labor share, it is useful to notice that equation (28) can be re-expressed as:

$$\int_0^{N_t} S_{i,t} a_{i,t} di = 1. \tag{30}$$

The term labeled *allocative efficiency* above would therefore equal 1 under a symmetric equilibrium, or in a model with additive aggregation of goods *and* a production function linear in labor. It follows that the term in bracket in equation (27) shows the contribution of the higher moments of the productivity and firm-size distributions to the aggregate productivity level.

**Aggregate Productivity Growth Rate**

I now shift the focus to the growth rate of aggregate productivity, which, together with the population growth rate, determines the growth rate of output per capita in the long-run.

**Proposition 1** *Under a time-invariant distribution of relative productivity levels, the long-run growth rate of aggregate productivity is a function of population growth and of the growth rate of the arithmetic average of firms' productivities.*

Proposition 1 highlights the sources of long-run growth. Its dependence on population growth is important in the context of studying the effects of slower labor force growth, which, as shown later, can explain the entire decline in the net entry rate. Its dependence only on the first moment of the aggregate productivity distribution ensures that firm-level productivity changes are the only relevant factors to consider in steady state. As long as the focus is on the steady state where the firm-size distribution is time-invariant, there is no concern over the aggregation of firm productivity increases.

The proposition can be expressed in a mathematical form starting from equation (27):

$$1+g^{productivity} = \underbrace{\left[\frac{n_t(z_t,\alpha_t)}{n_{t-1}(z_{t-1},\alpha_{t-1})}(1+\lambda)\right]^{\frac{1}{\epsilon-1}}}_{\text{semi-endogenous}} \underbrace{\left(1+g_t^A(z_{t-1},\alpha_{t-1},\alpha_t)\right)}_{\text{average productivity}} \underbrace{\left\{\frac{\frac{1}{N_t}\int_0^{N_t}\left[(S_{i,t}N_t)^\nu a_{i,t}\right]^{\frac{\epsilon-1}{\epsilon}}di}{\frac{1}{N_{t-1}}\int_0^{N_{t-1}}\left[(S_{i,t-1}N_{t-1})^\nu a_{i,t-1}\right]^{\frac{\epsilon-1}{\epsilon}}di}\right\}^{\frac{\epsilon}{\epsilon-1}}}_{\text{change in the distribution}}, \quad (31)$$

This expression resembles the one in Peretto and Connolly (2007), with the addition of the last term, which depends on heterogeneity in productivity levels and labor shares. The semi-endogenous component depends only on population growth in steady state as the average firm size is stationary. This term is sometimes referred to as *expanding variety*, and it emerges from the CES aggregator, which rewards a higher number of goods. $g_t^A$ is the growth rate of average productivity between $t-1$ and $t$, and it will be the focus of the remainder of the paper. Finally, the last term signals that aggregate productivity growth is dependent on changes in the distribution of relative productivity. Nevertheless, given a time-invariant distribution in steady state, the long-run growth rate of aggregate productivity is determined exclusively by the first two terms, while the last one is relevant along the transition, an exploration left for future research.

I now turn to showing the model's calibration before moving on to discuss the results of the paper.

## 2.5 Calibration

This subsection presents the calibration. The model is calibrated up to 1999, the period around which business dynamism starts changing more rapidly.

#### Externally Calibrated Parameters

I rely on external targets to match some parameters. These externally calibrated parameters are summarized in table 3. The labor force growth rate, $\lambda$, for the US from 1977 to 1999 is 1.8%. The parameter $\eta$ is set to 0.55 to match the return to labor in R&D in the US based on NSF data (Mand, 2019), while $\nu$, which determines the returns to labor in production, is set to 0.7, following Lee and

Mukoyama (2015), whose estimations of establishment-level productivity are also used to target the relative productivity of entrants of 0.75. $\beta = 0.995$ is selected to match, considering a growth rate of per capita consumption between 2.5% and 3%, a real rate of return of approximately 6.5% common in the literature (Akcigit and Ates, 2019). I further set $\epsilon = 6$ to match a markup over marginal cost of 20% and the elasticity of intertemporal substitution in consumption $1/\varphi = 0.5$, standard in the literature and in line with empirical estimates (Crump et al., 2015).

Hall et al. (2010), who review the literature on the returns to R&D, report a widely different ratio of social to private returns to R&D in the various estimations performed over the years. The only consensus seems to be that social returns are substantially larger than private returns. In line with Bloom et al. (2013), I target a ratio of social returns to private returns to knowledge of about 2.6, which requires $\mu = 0.28$.

Table 1: Externally calibrated parameters

| Parameter | Symbol | Target |
|---|---|---|
| Labor force growth | $\lambda$ | 1.8% |
| Returns to R&D labor | $\eta$ | 0.55% |
| Returns to production labor | $\nu$ | 0.7 |
| Discount rate | $\beta$ | real interest rate: 6.5% |
| Intertemporal elasticity of substitution | $1/\varphi$ | 0.5 |
| Elasticity of substitution between goods | $\epsilon$ | 20% markup |
| Relative knowledge of entrants | $\chi_z$ | relative productivity: 0.75 |
| Returns to knowledge | $\mu$ | Social to private returns to knowledge: 2.6 |

**Internally Calibrated Parameters**

The remaining parameters jointly determine the other targeted moments implied by the model. Some moments are particularly informative of the size of these internally calibrated parameters. The steady state entry and exit rates depend largely on the productivity of R&D of incumbents relative to entrants. $\chi_\alpha$ almost uniquely determines the average productivity growth of entrants, a moment that Huergo and Jaumandreu (2004) estimate to be 5% for Spanish manufacturing firms in selected industries in the period. $f_E$ is chosen to match the entry rate of new establishments, which, according to US Census data, averaged 13.4% between 1977 (the starting year of the series) and 1999. Moreover, the parameter $\sigma_\alpha^E$, which regulates the dispersion of the initial draw of the average productivity of R&D of entrants, is picked to match the average exit rate, 20.8%, of establishments younger than one year in the US in the same period.

Equation (23) shows that the parameter $\Phi$ affects the average size of establishments, which averages approximately 16.75 in the US from 1977 to 2014, implying $n_t = 0.0597$ in steady state.

Table 2: Internally calibrated parameters

| Parameter | Symbol | Value |
|---|---|---|
| Entrant's mean R&D productivity | $\log \chi_\alpha$ | -10.45 |
| Standard deviation entrants' R&D productivity | $\sigma_\alpha^E$ | 4.76% |
| Fixed operating cost | $\Phi$ | 5.1 |
| Fixed entry cost | $f_E$ | 2.72 |
| Shock persistence | $\rho$ | 0.721 |
| R&D productivity | $\log \overline{\alpha}$ | -10.45 |
| St. dev. relative knowledge startups | $\sigma_Z^E$ | 1.3 |
| St. dev. draw of R&D productivity shock | $\sigma_\xi$ | 2.901 |
| Returns to knowledge in production | $\theta$ | 0.1259 |

The parameters $\rho$ and $\sigma_\alpha$ in addition to $\overline{\alpha}$ affect the average productivity of R&D of incumbents in the economy. I target an average productivity growth rate of establishments of 3%, in line with the aforementioned estimate of Huergo and Jaumandreu (2004). Although this moment is computed from a sample of Spanish firms in selected R&D performing industries in the '90s, the aggregate growth rate implied by the model calibrated this way is very similar to the one of the IT sector in the US in the same period shown in Byrne et al. (2016), and also to the one of US R&D performing industries. At the same time, the combination of these three parameters regulates the extent to which productivity improvements are expected or unexpected at the time of undertaking the investment, and the persistence of sales growth rates. I rely on Doraszelski and Jaumandreu (2013), who estimate the ratio of the variance of the unexpected component of productivity change and the variance of the change in productivity for the same sample of Spanish manufacturing firms. As their estimates vary by industry ranging from 25% to 75%, I use a value of 31%, which is one standard deviation below the average across those industries to be conservative. Given this variability across industries, I choose to be conservative in this case, to avoid overattributing job reallocation to the process of technological knowledge creation. The target for the average autocorrelation of sales growth rates is instead 0.2, which is the post-war average for the US economy (Dosi et al., 2019). The standard deviation of relative knowledge for entrants is calibrated to match the ratio between the standard deviation of employment for startups and their standard deviation the following year. In the US economy, this value is 1.02 (Sterk et al., 2021).

Finally, the parameter $\theta$ determines the degree of knowledge intensity of economic activity, affecting the returns to R&D. I use the R&D to GDP ratio, which is 2.4% up to the late '90s.[8]

Table 3: Targets for internally calibrated parameters

| Parameter | Source | Value |
|---|---|---|
| Average plant size | BDS | 16.75 |
| R&D to GDP | NIPA | 2.4% |
| Entry rate | BDS | 13.4% |
| Startup's exit rate | BDS | 20.8% |
| Unexpected vs expected $\Delta A_i$ | Doraszelski and Jaumandreu (2013) | 31% |
| Average growth rate | Huergo and Jaumandreu (2004) | 3% |
| Persistence growth rates | Dosi et al. (2019) | 0.2 |
| St. dev. employment 2-years to 1-year firms | Sterk et al. (2021) | 1.02% |
| Average growth entrants | Huergo and Jaumandreu (2004) | 5% |

**Non-targeted Moments**

Table 4 summarizes some relevant non-targeted moments. These moments are the ones that will be discussed in the quantitative exercise about causes and consequences of changing business dynamism.

I compare the labor productivity growth rate to the one for the US in R&D performing industries during the period '91-'99, which coincides to the period considered for the targeted moments on firm-level growth.[9] This calibration is the only one in the literature that does not use a measure of aggregate growth as a target. However, this value reproduced by the model is very close to the analogous one observed in the data.

Job reallocation has potentially different causes, among which the one explored in this paper, namely, variability in firms' R&D investment decisions and outcomes. Growth rates exhibit long

[8]In the model, the R&D to GDP ratio requires a definition of GDP. I obtain it starting from equation (2):

$$\underbrace{w_t \int_0^{N_t} l_{Z_{i,t}} di}_{\text{R\&D}} + \underbrace{c_t L_t}_{\text{Consumption}} + \underbrace{Z_t^{\theta} f_E N_{E_t}}_{\text{Investment}} = \underbrace{\int_0^{N_t} \pi_{i,t} di + w_t \int_0^{N_t} l_{Z_{i,t}} di}_{\text{Operating Profit}} + \underbrace{w_t L_t}_{\text{Labor}} \equiv GDP_t. \tag{32}$$

[9]Labor productivity is obtained in the data by dividing the total of net sales in R&D performing industries to the total of employment in the same industries, and the GDP deflator.

left and right tails. In this model, only the right tail is present because employment depends exclusively on productivity and the maximum rate at which firms can lose productivity is the aggregate growth rate since no knowledge regress is conceived. As differences in R&D investment and outcome are only one of the factors that drives job reallocation, looking at this non-targeted moment can shed light on how quantitatively important this process is. Additionally, the conservative calibration of the R&D productivity shock's variance increases confidence that job reallocation is at least not overestimated. The model explains slighly above half of the job reallocation rate observed in the US data in the period 1977 to 1999. The takeaway is that differences in the ability to innovate are an important driver of job reallocation across firms.

When it comes to the process of entry and exit, the model performs well in terms of productivity, but less well in terms of employment. The relative productivity of entrants is exact because it is a targeted moment, while the relative productivity of exiters is almost the same as it is overestimated by only 3%. The relative employment levels of entrants and exiters are overestimated respectively by approximately 30% and 15%. This can introduce a bias on the reallocation rate, and indeed this is what happens. For this reason, in the quantitative exercise I focus on matching changes in the reallocation rate across incumbents rather than on the reallocation rate for the whole economy.

Finally, the firms' growth rates are negatively correlated with size, in line with the data. A quantitative estimation of this and a comparison with empirical estimates is offered in appendix C.

Table 4: Relevant Non-targeted Moments

| Moment | Model | Data | |
|---|---|---|---|
| Reallocation rate | 28% | 30.6% | US Census |
| Reallocation rate (cont.) | 10.5% | 18.4% | US Census |
| Share Reall. explained (cont.) | 57% | 100% | US Census |
| Lab. productivity growth | 3.8% | 3.6% | NSF |
| Mean productivity exit | 0.66 | 0.64 | Lee and Mukoyama (2015) |
| Rel. emp. entrants | 0.77 | 0.6 | Lee and Mukoyama (2015) |
| Rel. emp. exiters | 0.57 | 0.49 | Lee and Mukoyama (2015) |

# 3 Sources of Firm Growth, Churning, and Stationarity

As discussed in Proposition 1, the economy's growth rate depends on changes in firms' productivity levels. Moreover, focusing on firm growth is relevant to understanding what causes churning and drives the firm-size distribution's stationarity. I illustrate the exact relation between firm growth rates and aggregate growth rates in the appendix.

This section is devoted to showing the sources of churning and the stationarity of the distribution of firms over the state variable, relative knowledge, and the exogenous shock, R&D productivity. Furthermore, it illustrates how life cycle dynamics arise.

Churning arises whenever there are differences in growth rates across firms, which lead to movements within the relative productivity and firm-size distributions. These movements are due partly to the realization of the shock and partly to firms' decision rules. The first subsection covers this in detail.

The existence of a non-degenerate, stationary, ergodic distribution is a related issue. The exogenous shock evolves according to a stationary process by assumption. Nevertheless, one would need to verify that differences in growth rates are not such that larger firms grow faster than smaller ones, thus manifesting a tendency towards monopoly.

The first subsection discusses the forces that drive differences in growth rates across firms of different relative sizes. The second subsection presents the phase diagram that illustrates the convergence process of each firm. The third subsection illustrates the firm-size distribution generated by the model.

## 3.1 Growth Rate Differentials

In what follows, I show the sources of growth rate differentials across firms, which cause churning. Additionally, I show that expected growth rates are strictly decreasing in relative knowledge conditional on R&D productivity, which drives the stationarity and non-degeneracy of the distribution.

By using the approximation $g_{t+1}^{A_i} \approx \theta g_{t+1}^{Z_i}$, I can derive the ratio of growth rates of two arbitrary firms from equation (12) after plugging in the optimal $l_Z$ value:

$$\frac{g_{t+1}^{A_i}}{g_{t+1}^{A_j}} \approx \frac{\alpha_{i,t+1} z_{i,t}^{\mu-1}}{\alpha_{j,t+1} z_{j,t}^{\mu-1}} \left[ \frac{l_{Z_{i,t}}(z_{i,t}, \mathbb{E}_t \alpha_{i,t+1})}{l_{Z_{j,t}}(z_{j,t}, \mathbb{E}_t \alpha_{j,t+1})} \right]^{\zeta}. \tag{33}$$

This ratio is affected by three elements: the realized R&D productivity, the initial relative knowledge level, and the R&D effort exerted by the firm. The first obvious source of churning is the realization of the R&D productivity shock. I refer to this as *unsystematic churning*. Note that all the remaining sources of churning create differences in the expected productivity growth rates. I refer to differences in expected productivity growth rates as *systematic churning*. As the R&D effort is a function of relative knowledge and expected R&D productivity, the expectation of the growth rate differential at time $t+1$ depends exclusively on the state variable and the exogenous shock. A variety of elements affect this expectational growth rate differential. I discuss them in sequence.

First, the term $z^{\mu-1}$ suggests that for any given R&D investment and expectation of R&D productivity, the growth rate declines in relative knowledge, as the private returns to knowledge $\mu < 1$. This effect is driven by the knowledge spillover, which operates as a force of attraction: firms above the average knowledge level will be dragged down in relative terms by the spillover, while firms below the average knowledge level will be lifted up by it. The larger the firm, the larger the R&D productivity and investment required to balance this force of attraction. The presence of private returns to knowledge partially offsets this effect by facilitating the accumulation of knowledge for those firms that already possess more of it.

Second, R&D is a function of relative knowledge. A necessary condition to deliver this dependence is that the innovator must know in advance what product-line they will improve when taking the investment decision. However, this condition is absent in several growth models. In this framework, firms perform R&D in-house, in line with empirical evidence (Dosi, 1988). Therefore, firms consider their relative productivity level — and consequently, their market share — when investing.

R&D's dependence on relative knowledge and R&D productivity can be analyzed by considering the ratio of R&D investment for two arbitrary firms. Differentiating the detrended Bellman equation with respect to $l_{Z_{i,t}}$ for firms that do not exit the market at time $t$, one can derive an

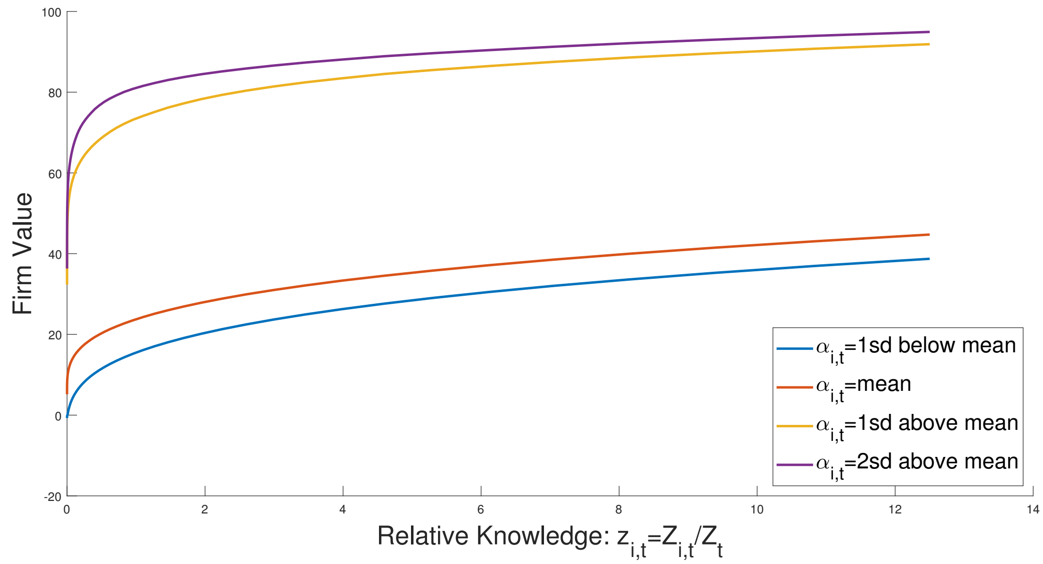


Figure 1: Value functions for selected R&D productivity levels.

expression for this ratio:

$$\frac{l_{Z_{i,t}}}{l_{Z_{j,t}}} = \left[ \frac{\mathbb{E}_t[\alpha_{i,t+1}] z_{i,t}^{\mu} \frac{\partial \mathbb{E}_t V(z_{i,t+1}, \alpha_{i,t+1})}{\partial \mathbb{E}_t z_{i,t+1}}}{\mathbb{E}_t[\alpha_{j,t+1}] z_{j,t}^{\mu} \frac{\partial \mathbb{E}_t V(z_{j,t+1}, \alpha_{j,t+1})}{\partial \mathbb{E}_t z_{j,t+1}}} \right]^{\frac{1}{1-\zeta}}. \tag{34}$$

As equation (34) shows, R&D effort is not equal across firms. Conditional on relative knowledge, firms with a higher expected R&D productivity will invest more. The dependence on relative knowledge is instead a more complex one. Relative knowledge shows up due to two forces: the ability to internalize their knowledge in R&D and the gain in expected value that more knowledge can achieve.

Figure 1 shows the shape of the value function for selected R&D productivity levels, which is increasing at a diminishing rate. The slope of the value function declines in relative knowledge even though both the production and the R&D technology allow for increasing returns in the private factors of production. Importantly, diminishing returns to size originate from the demand side through a mechanism that resembles the one emphasized in a different context by Acemoglu and Ventura (2002). Firms that gain more technological knowledge relative to others increase their volume of production. By producing more, they however face a lower price as product differentiation ensures that firms face a downward sloping demand curve. This price reduction is in turn responsible for dragging down the return to further knowledge accumulation. As a result, incentives to innovate decline as firms grow larger relative to others.

The concavity of the vale function has an additional effect since the incentive to innovate is a function of the slope of the *expected* value function. This concavity creates a Jensen's inequality between the expected value of the product-line and the value of the expectation. Jensen's inequality is stronger where the curvature is more pronounced, i.e., for small firms and firms with a high R&D productivity since they can expect a larger change in knowledge. Instead, it weakens as the curve flattens out.

Finally, private returns to knowledge increase the incentive to innovate for firms that possess more of it.

The three listed effects deliver a dependence of R&D investment on relative productivity, as illustrated in figure 7 in the appendix. As relative knowledge (therefore productivity) tends to infinity, the value function flattens, and the only relevant effect is the direct effect of diminishing returns, implying that R&D tends to 0.

As shown in figure 2, larger firms grow in expectation slower everywhere, conditional on R&D productivity. The R&D investment differential across firms with the same expected R&D productivity reinforces the growth rate differential for large firms while having an ambiguous effect on small ones.

The following subsection explores the convergence process in further detail by considering the state variable and the exogenous shock jointly and focusing on entry and exit.

### 3.2 Mobility and Stationarity

I now show the individual firm's expected convergence process that gives rise to the stationary distribution over the state variable and the exogenous shock. I also show the distribution of firms over the state-space. I will finally make some more general considerations about the nature of aggregate productivity growth as a process influenced by the product life cycle.

Figure 3 illustrates the phase diagram that describes this convergence proces by showing the expected evolution of firms over relative productivity and R&D productivity. The LL locus shows the long-run expectation of the exogenous AR(1) process that characterizes the evolution of R&D

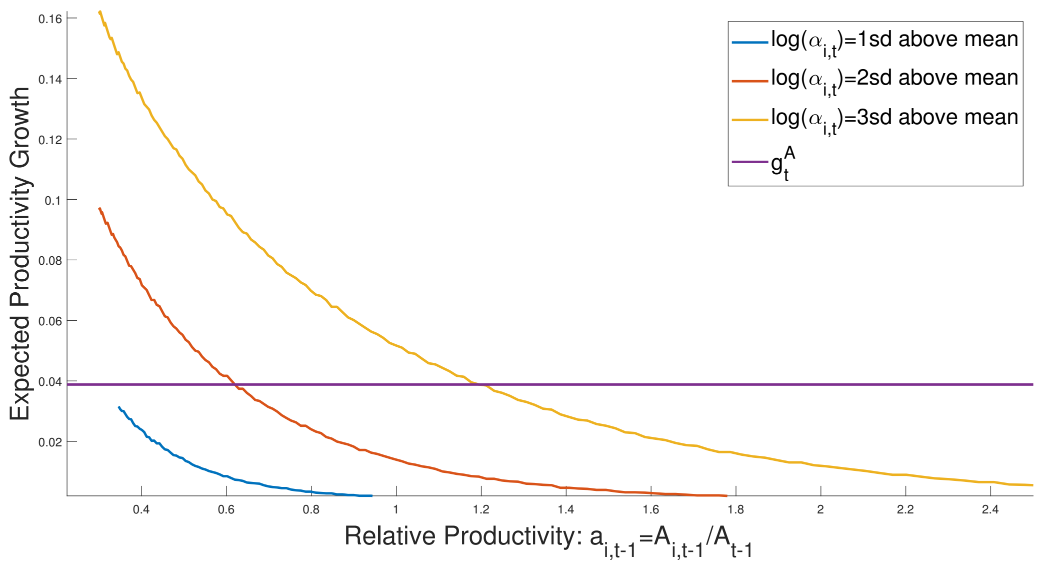


Figure 2: Expected growth rates of productivity and relative productivity levels for selected levels of R&D productivity, 1, 2, and 3 standard deviations above the mean. The horizontal line is the growth rate of aggregate productivity.

productivity.

The convergence process over relative productivity can be understood by analyzing the R&D technology in realization for surviving firms given by equation (12), combined with the policy function, which can be rearranged to yield:

$$\frac{a_{i,t+1}(z_{i,t+1})}{a_{i,t}(z_{i,t})} = \frac{(1+\alpha_{i,t+1} z_{i,t}^{\mu-1} l_{Z_{i,t}}(\alpha_{i,t}, \mathbb{E}_t \alpha_{i,t+1}, z_{i,t})^{\zeta})^{\theta}}{1+g_{t+1}^A}. \tag{35}$$

At this point, it is possible to construct a locus over $a_{i,t}$, $\alpha_{i,t+1}$, and $\alpha_{i,t}$, along which the relative productivity level remains constant over time. I call this the *no-churning locus*. Nevertheless, it is more useful to define the no-churning locus in only two dimensions, namely the combination of $a_{i,t}$, and $\alpha_{i,t}$ for which relative productivity remains *on average*, or *in expectation*, constant over time. That implies a no-churning locus given by:

$$1+g_{t+1}^A = \mathbb{E}_t (1+\alpha_{i,t+1} z_{i,t}^{\mu-1} l_{Z_{i,t}}(\alpha_{i,t}, \mathbb{E}_t \alpha_{i,t+1}, z_{i,t})^{\zeta})^{\theta}. \tag{36}$$

For any R&D productivity level, and conditional on survival, convergence to the no-churning locus requires firms' growth rates to decline in relative productivity. Figure 2 from the previous subsection shows the relationship between surviving firms' average growth rate for selected initial R&D productivity levels and relative productivity. The no-churning locus shows the values of relative productivity and R&D productivity at which growth rates are equal across firms.

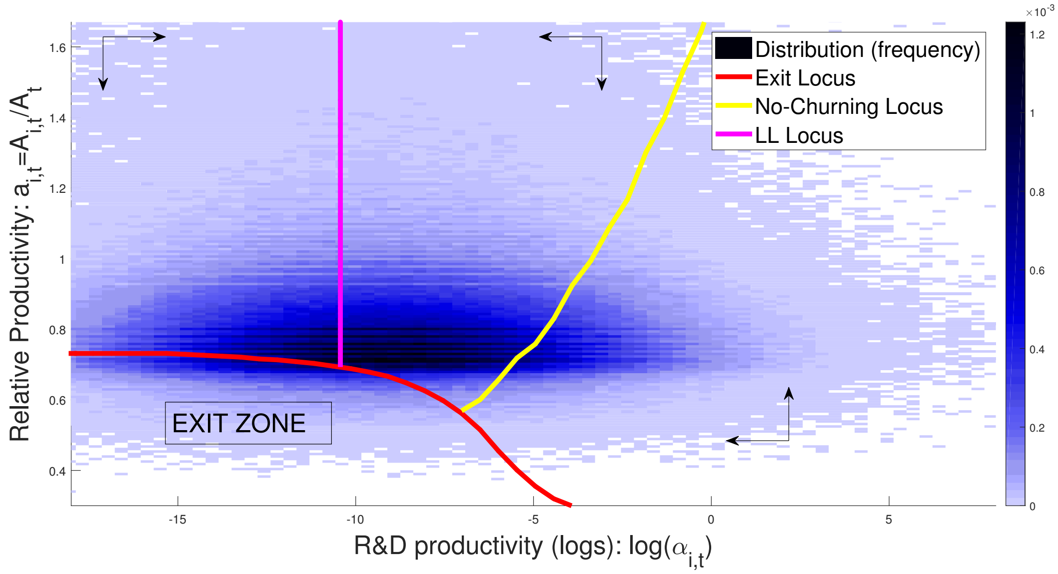


Figure 3: Phase diagram and firm density over the state variable and the exogenous shock. The no-churning locus represents the combination of relative productivity and R&D productivity for which the firm's expected growth rate equals the growth rate of average productivity. The LL locus shows the long-run expectation of R&D productivity. The exit locus shows the combination of relative productivity and R&D productivity below which the firm exits the market. The figure also shows the distribution of firms simulated using the model as the data generating process.

As the problem is stochastic, firms are never on the no-churning locus. Therefore, growth rates are not equalized at each point in time, but only on average. Firms below it will grow at a faster rate than average productivity and vice-versa. Furthermore, as figure 3 shows, the no-churning locus is upward sloping, implying that a higher relative productivity level requires a higher R&D productivity to avoid shrinking. This positive slope illustrates that firms with a persistently higher ability to innovate will eventually manifest it in their size and not in their growth rate.

**Proposition 2** *If productivity growth rates conditional on the R&D productivity level exhibit negative scale dependence and if the idiosyncratic R&D productivity level evolves stochastically, at least some firms are away from their long-run relative productivity level at any time, generating churning endogenously in the form of mean-reversion.*

Proposition 2 expresses that, although firms tend endogenously towards the no-churning locus, the shock disrupts their position in the state-space in every period. The resulting systematic churning is a manifestation of mean reversion. This is one of the key results of the paper: churning, hence turbulence, arises endogenously as the result of firms' optimization.

Aggregate productivity growth and turbulence are manifestations of the same process: firms' competing for market share to maximize their value. Because market share is gained by improving technological knowledge through R&D, aggregate productivity increases. Because the marginal value of gains in market share varies depending on current market share, incentives to perform R&D change with relative size, giving rise to turbulence.

Turbulence is driven also by entry and exit. The following subsection explores these processes.

#### 3.2.1 Entry and Exit: The Firm Life-Cycle and Gradual Creative Destruction

The presence of entry and exit does not disrupt the stationarity of the distribution, but it introduces selection effects. Indeed, the process of entry and exit is the one described in Hopenhayn (1992). The crucial difference with Hopenhayn's model is that, in this case, firm productivity evolves endogenously. In what follows, I show that selection has two relevant implications: first, it endogenizes the cross-sectional average of the exogenous process as firms with low levels of R&D productivity exit the market. Second, it introduces an expected life trajectory for each firm.

Figure 3 includes an exit locus. This locus is an absorbing barrier: firms whose levels of relative productivity and R&D productivity lie below that curve have a negative continuation value and exit at the end of the period. Unlike models of firm dynamics, as the firm's value depends on the endogenous state variable, the distribution over productivity does not have an abrupt truncation but a smoother left tail.

The first first implication of endogenous exit emerges from the same figure, which shows the firms' distribution over relative productivity and R&D productivity. R&D productivity's cross-sectional average is higher than its unconditional expectation depicted on the LL locus. This inequality arises because exiting firms have, on average, a lower R&D productivity. Furthermore, with this calibration, entrants have a higher R&D productivity than incumbents. Thus entry increases the R&D productivity's cross-sectional average even further.

In turn, the disconnect between the R&D productivity cross-sectional average and the individual firm's time average implies that each firm has an expected finite life. Figure 4 shows the expected life trajectory of three selected entrants that differ in their initial R&D productivity draw. All firms expect to converge to the exit locus eventually.

These dynamics highlight the relevance of the firm life-cycle for the growth process. Furthermore, they illustrate the nature of creative destruction as a gradual process. Given the calibration presented earlier, surviving entrants tend to have a higher ability to innovate than incumbents. They do not directly replace incumbents, but they gain relative productivity over time, stealing their market share. As a result, less innovative firms keep losing ground until they exit the market when their relative productivity level is low enough to make them unprofitable.

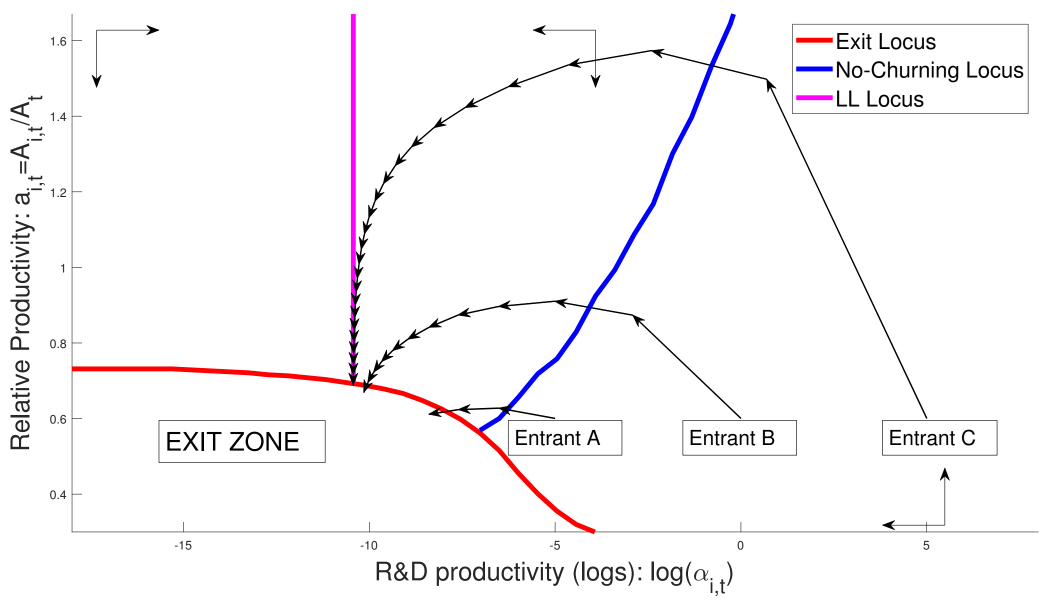


Figure 4: Phase diagram over the state variable and the exogenous shock. The no-churning locus represents the combination of relative productivity and R&D productivity for which the firm's expected growth rate equals the growth rate of average productivity. The LL locus shows the long-run expectation of R&D productivity. The exit locus shows the combination of relative productivity and R&D productivity below which the firm exits the market. The figure also shows the expected life path of three startups that differ in their initial R&D productivity draw.

### 3.2.2 Firm-Size Distribution

A contribution of this paper is to describe the forces that reproduce endogenously the shape of the firm-size distribution observed in the data.

Figure 5 plots the firm-size distribution in logarithmic scale compared to a Pareto distribution.

The distribution shape resembles the one observed in the data shown in figure 1 of Rossi-Hansberg and Wright (2007). In particular, the firm size distribution resembles a Pareto distribution except for very small and very large firms, the left and right tails, respectively.

Deviations from a Pareto distribution on the right are the automatic consequence of negative scale dependence (Rossi-Hansberg and Wright, 2007). The most successful firms become very large through a series of positive shocks and R&D. Eventually, their R&D would decline, tending towards zero due to the mechanism that drives scale dependence illustrated above. As R&D tends towards zero, the probability that these firms can maintain their size tends to zero while the rest of the economy grows.

Deviations from a Pareto distribution on the left are the result of small firms exiting the market.

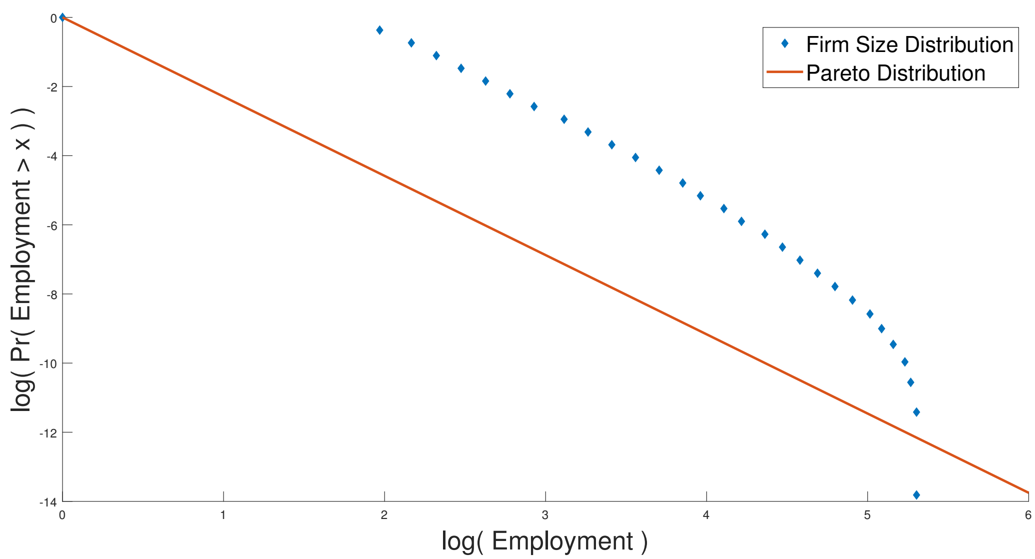


Figure 5: Comparison between a counter cumulative distribution function (in logarithmic scale) for the firm sizes generated by the model and for a fitted Pareto distribution. The observations generated by the model are aggregated into 30 bins.

## 4 Consequences of Changing Business Dynamism

What consequences do changes in business dynamism entail for aggregate productivity growth? This section tackles this question. Business dynamism has been changing in the past few decades in the US, with strong discontinuities in job reallocation and entry/exit around the year 2000 (Decker et al., 2016, 2020). Aggregate productivity growth has instead transitioned from high to low levels twice in the post-WWII period. As commonly broken down in the literature (Syverson, 2017), labor productivity growth averaged 2.7% from 1947 to 1973, before falling by approximately half to 1.5% in 1974-1994. It again rose to 2.8% from 1995 to 2004, followed by a more moderate 1.3% from 2005 to 2015.

Two sets of yet unexplained facts emerge from this picture: first, is the reduction in productivity growth from the period '94-'04 to the period '05-15 imputable to the same causes of changing business dynamism? The debate on this is ongoing, and the question is still unresolved (Naude, 2020).

Second, the sample is characterized by an increase in the R&D to GDP ratio that has not led to faster growth. What can reconcile this disconnect between R&D and productivity growth? A hypothesis for this phenomenon is that ideas are getting harder to find, namely, that the average R&D productivity is on a decline implying that more R&D is needed to sustain the same productivity growth rate (Bloom et al., 2020).

The section presents a counterfactual analysis through a few quantitative exercises to illustrate the effect of changes in selected parameters on the dimensions of business dynamism and aggregate productivity growth. I compare a steady state calibrated as before on the period '77-'99 to another steady state calibrated in the same way, but with parameter differences that lead to changes in the dimensions of business dynamism averaged over the period 2000-2014.[10] I focus on changes in labor force growth (in appendix), entry costs, fixed operating costs (in appendix), and knowledge diffusion. As discussed in each subsection, a large literature documents changes in these factors and links them to changes in specific dimensions of business dynamism. I show that these changes in business dynamism can be reproduced here as well. Then, I illustrate the qualitative effects of these parameters changes, I quantify their effects in isolation, and finally, I combine all changes to draw conclusions on aggregate productivity growth and R&D.

### 4.1 Increase in Entry Cost

Table 5 shows the change of selected moments following a 32% increase in the entry cost, which matches the change in entry rate observed in the data when comparing the periods '77-'99 to

[10] Using 2004 instead of 2000 as the threshold year shows a slightly stronger reduction of aggregate productivity growth, but it does not change the conclusions.

’00-’14. Evidence of increasing entry costs is presented in Gutiérrez and Philippon (2019).

Table 5: Entry cost increases by 32% to match the change in entry

| Moment | Model | Data | Source |
|---|---|---|---|
| Δ Net entry rate | 0pp | -1pp | US Census |
| Δ Entry rate (targeted) | -2.4pp | -2.4pp | US Census |
| Δ Reallocation rate | -4.3pp | -4pp | US Census |
| Δ Reallocation rate (cont.) | +0.1pp | -0.7pp | US Census |
| %Δ Lab. productivity growth | -1% | -53% (-1.5pp) | FRB SF (’94-’04 vs ’05-’15 ) |
| | | -14% (-0.2pp) | FRB SF (’73-’93 vs ’05-’15 ) |
| Δ R&D to GDP | +15% (+0.35pp) | +9%(+0.2pp) | NIPA (’94-’04 vs ’05-’15 ) |
| | | +23% (+0.4pp) | NIPA (’77-’99 vs ’00-’14) |
| %Δ Average plant size | +1% | +4.4% | US Census |

By construction, the change in entry cost has a direct negative effect on the entry rate. Additionally, this has interesting implications.

The effect on aggregate productivity growth is the combination of countering forces. First, an increase in the entry cost reduces the aggregate growth rate, although slightly. This outcome is in contrast with Peretto and Connolly (2007), a version of this model with a symmetric equilibrium, where a lower entry rate reduces the number of firms present in the market, thus increasing the average expected market share and, as a consequence, the returns to R&D. Although this cost-spreading effect is present in this framework too, it is not the only force at play. The selection effect emphasized above is responsible for two occurrences: reducing the number of entrants, which on average are more innovative than incumbents, and shielding low R&D productivity firms from entrants’ competition, thus encouraging them to stay in the market. As a consequence, the cross-sectional average of the R&D productivity decreases.

Additionally, although aggregate productivity growth drops, R&D as a share of GDP increases, despite the reduction in average R&D productivity mentioned above. This trend is visible in the data (Bloom et al., 2020), and the model can rationalize it without attributing it to any change in the parameters of the R&D technology. The forces responsible for this result are: first, the larger average market share fuels a stronger cost spreading effect. Second, small firms that would otherwise exit remain in the market. This change in the firms’ population strengthens the *scale dependence effect*: a smaller relative size provides stronger incentives to innovate, all else constant. These small firms have a higher incentive to invest than equally innovative but larger firms because they are currently charging a higher price, which gives them a higher return to increasing their relative size.

### 4.2 Increase in Private Returns to Knowledge

Here, I consider an increase in the private returns to knowledge (and a consequent reduction in the social returns) by 11.5% to match the change in the exit rate of 1-year old establishments. There are various indications that knowledge diffusion has decreased in the past decades, mainly due to patent protection (Akcigit and Ates, 2019) and changes in innovation policy (Belenzon and Cioaca, 2021).

Furthermore, Akcigit and Ates (2019) identified this channel as the one that can qualitatively match all changes in business dynamism in their model. My result differs from theirs in this respect because I allow for exit as an optimal stopping problem and include a free entry condition. In my framework, reducing knowledge diffusion has two significant partial equilibrium effects on both entry and exit.

First, given a realistic parametrization according to which entrants expect to start with a below-average knowledge level, entry is discouraged because this change hinders their potential to grow.

Second, exit is encouraged because the force of attraction that can lift low-knowledge firms is weaker.

The overall effect on the steady state gross entry rate depends on which effect prevails. In this case, a large share of firms clusters around the exit locus. Therefore, the effect on exit is stronger and more firms leaving the market causes more firms to enter through general equilibrium effects, increasing the overall number of firms.

Churning and, therefore, reallocation is affected as well. According to equation (34), knowledge diffusion regulates the force of attraction towards the average knowledge level. Therefore, its reduction directly affects the growth rate of low-knowledge firms negatively and positively the growth of high-knowledge firms. Consequently, expected growth rates are closer together, reducing churning.

A further partial equilibrium effect is the one that runs through the investment decisions of firms. As private returns to knowledge increase, the incentive to innovate increases too. Simultaneously, social returns to knowledge are lower. As a result, this effect is not uniform for all firms, as those with a knowledge level higher than average would see their knowledge change per unit of R&D be higher for higher $\mu$, while the opposite is true for low knowledge firms, whose R&D may even decrease. Additionally, general equilibrium effects modify firms' decisions through the change in the average firm size, which decreases, implying a weaker cost-spreading effect. Therefore, the reduction in knowledge diffusion may fuel faster or slower unsystematic churning that may counter or reinforce the effect of weaker systematic churning. Overall, the quantitative analysis shows that the effects that hinder reallocation prevail.

Finally, aggregate productivity growth is itself affected by multiple countering forces. The weaker force of attraction is detrimental for growth because the distribution is skewed, and most firms possess a below-average knowledge level. Additionally, the effects on entry and exit could determine a faster or slower product life cycle. Finally, there are multiple effects on the incentive to innovate, as mentioned above. Overall, in this case, growth is reduced slightly.

Table 6: Private returns to knowledge $\mu$ increase by 11.5% to match the change in exit rate of one-year-old plants

| Moment | Model | Data | Source |
|---|---|---|---|
| Δ Net entry rate | 0pp | -1pp | US Census |
| Δ Entry rate | +0.7pp | -2.4pp | US Census |
| Δ Reallocation rate | +0.7pp | -4pp | US Census |
| Δ Reallocation rate (cont.) | -0.1pp | -0.7pp | US Census |
| %Δ Lab. productivity growth | -1% | -53% (-1.5pp) | FRB SF ('94-'04 vs '05-'15 ) |
| | | -14% (-0.2pp) | FRB SF ('73-'93 vs '05-'15 ) |
| Δ R&D to GDP | +6% (+0.15pp) | +9%(+0.2pp) | NIPA ('94-'04 vs '05-'15 ) |
| | | +23% (+0.4pp) | NIPA ('77-'99 vs '00-'14) |
| Δ Exit rate 1-year (targeted) | +1.1pp | +1.1pp | US Census |
| %Δ Average plant size | -1.2% | +4.4% | US Census |

### 4.3 Declining Business Dynamism: Combining All Changes

Table 7 shows the effect of simultaneous changes on labor force growth, knowledge diffusion, entry cost, and fixed operating cost to match changes in reallocation across continuing incumbents, entry rate, and average establishment size in the US economy between the periods '77-'99 and '00-14. In this way, I provide a quantitative estimate of changing business dynamism's effects on aggregate productivity growth.

The channels through which aggregate productivity growth is affected are all those mentioned in the previous subsections. However, the quantitative magnitude of those effects may be different when all changes co-occur. For instance, as the increase in the fixed entry cost pushes the exit locus downwards, fewer firms are clustered around it, and the effect on the gross entry rate of weaker knowledge diffusion changes sign. This reversal has implications for growth since the product life cycle effect is slower, and the average R&D productivity declines.

Table 7: Labor force $\lambda$ down by 1pp. $\Phi$ up by 7%, $\mu$ up by 33%, $F_E$ up by 36% matching average size, incumbents reallocation, and entry respectively.

| Moment | Model | Data | Source |
|---|---|---|---|
| Δ Net entry rate | -1pp | -1pp | US Census |
| Δ Entry rate (targeted) | -2.4pp | -2.4pp | US Census |
| Δ Reallocation rate | -3.4pp | -4pp | US Census |
| Δ Reallocation rate (cont.) (tar.) | -0.7pp | -0.7pp | US Census |
| %Δ Lab. productivity growth | -13% | -53% (-1.5pp) | FRB SF ('94-'04 vs '05-'15 ) |
| | | -14% (-0.2pp) | FRB SF ('73-'93 vs '05-'15 ) |
| Δ R&D to GDP | +26% (+0.6pp) | +9%(+0.2pp) | NIPA ('94-'04 vs '05-'15 ) |
| | | +23% (+0.4pp) | NIPA ('77-'99 vs '00-'14) |
| %Δ Average plant size (tar.) | +4.4% | +4.4% | US Census |

A remarkable result of this analysis is that the reduction in aggregate productivity growth is accompanied by an increase in R&D. This disconnect between the two is visible in the data. It has been interpreted, for example, by Bloom et al. (2020) as an indication that ideas are getting harder to find. This model links this phenomenon to parameters that explain changing business dynamism. The result is only partially due to changes in the average R&D productivity in the economy that arise endogenously due to selection effects. As explained above in the context of increasing entry costs, R&D increases because of stronger scale dependence effect and cost spreading effect.

Although the model delivers a reduction in aggregate productivity growth associated with declining business dynamism, the quantitative effect is not large enough to explain the productivity slowdown relative to the '94-'04 period. This counterfactual analysis brings a new argument to the debate over the productivity slowdown, suggesting that the changes that drove the decline in business dynamism could not have caused the high-growth regime of the mid-'90s to mid-2000s.

Additionally, these parameter changes lead to a reallocation of R&D efforts across firms. By keeping alive more low-knowledge and low R&D productivity firms, thanks to a higher entry cost, more R&D will come from them. As explained above, this occurs because their incentive to invest increases as they charge a relatively high price, but the contribution to expanding the average stock of knowledge will be small due to their low ability to innovate.

# 5 Conclusion

This paper presents a unified framework for studying aggregate productivity growth and business dynamism in all its dimensions. The model features a continuum of monopolistically competitive firms subject to idiosyncratic shocks to their R&D productivity. Due to in-house R&D and diminishing returns per good, R&D investment ultimately declines in relative size, causing churning to arise endogenously. Therefore, aggregate productivity growth is a turbulent process characterized by movements within the firm-size distribution as firms adjust their size optimally by choosing their innovation effort. Entry and exit create life-cycle effects that foster this systematic churning and shape the aggregate growth process through *gradual creative destruction*.

The paper introduces a couple of other theoretical contributions. First, it merges and advances two independent streams of literature, on firm dynamics (Hopenhayn, 1992) and on endogenous growth (Peretto, 1998; Dinopoulos and Thompson, 1998; Young, 1998; Peretto and Connolly, 2007), that are complementary to each other. I add endogenous productivity growth to the former, and I add turbulence to the latter.

Second, it identifies in the R&D scale dependence a novel mechanism that delivers a stationary firm size distribution whose right tail resembles the one observed empirically (Rossi-Hansberg and Wright, 2007).

Finally, I perform a counterfactual analysis to shed light on the consequences of changing business dynamism, obtaining two main results. First, the model suggests that the plausible causes of changing business dynamism do not explain the large decline in aggregate productivity growth observed when comparing 1994-2004 and 2005-2015 in the US. Although the causes of changing business dynamism lead to the decline of aggregate productivity growth, this drop is much smaller than the one observed in the data. This result contributes to the debate on the causes of the productivity slowdown summarized in Naude (2020).

Second, the model can rationalize both from a qualitative and quantitative perspective why R&D has increased without a proportional increase in the productivity growth rate. The causes of changing business dynamism are responsible for a stronger cost-spreading effect and keeping alive firms with a low ability to innovate but a relatively large incentive to do so due to their small market share. At the same time, they affect the characteristics of the population of firms in the market, reducing the average R&D productivity, with negative consequences for productivity growth.

## References


Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, Reallocation, and Growth. *American Economic Review*, 108(11):3450–3491.

Acemoglu, D. and Cao, D. (2015). Innovation by entrants and incumbents. *Journal of Economic Theory*, 157:255–294.

Acemoglu, D. and Ventura, J. (2002). The World Income Distribution. *The Quarterly Journal of Economics*, 117(2):659–694.

Akcigit, U. and Ates, S. T. (2019). What Happened to U.S. Business Dynamism? Working Paper 25756, National Bureau of Economic Research.

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

Audretsch, D., Klomp, L., Santarelli, E., and Thurik, A. (2004). Gibrat's Law: Are the Services Different? *Review of Industrial Organization*, 24(3):301–324.

Audretsch, D. B., Santarelli, E., and Vivarelli, M. (1999). Start-up size and industrial dynamics: Some evidence from Italian manufacturing. *International Journal of Industrial Organization*, 17(7):965–983.

Belenzon, S. and Cioaca, L. C. (2021). Guaranteed Markets and Corporate Scientific Research. Technical Report w28644, National Bureau of Economic Research.

Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2012). Endogenous Entry, Product Variety, and Business Cycles. *Journal of Political Economy*, 120(2):304–345.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are Ideas Getting Harder to Find? *American Economic Review*, 110(4):1104–1144.

Bloom, N., Schankerman, M., and Reenen, J. V. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 81(4):1347–1393.

Bottazzi, G., Cefis, E., Dosi, G., and Secchi, A. (2007). Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries. *Small Business Economics*, 29(1-2):137–159.

Brown, C., Haltiwanger, J., and Lane, J. (2008). *Economic Turbulence: Is a Volatile Economy Good for America?* University of Chicago Press.

Byrne, D. M., Fernald, J. G., and Reinsdorf, M. B. (2016). Does the United States Have a Productivity Slowdown or a Measurement Problem? *Brookings Papers on Economic Activity*, 2016(1):109–182.

Cao, D., Hyatt, H. R., Mukoyama, T., and Sager, E. (2017). Firm Growth through New Establishments.

Caves, R. E. (1998). Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature*, 36(4):1947–1982.

Crump, R. K., Eusepi, S., Tambalotti, A., and Topa, G. (2015). Subjective Intertemporal Substitution. Federal Reserve Bank of New York Staff Reports 734, Federal Reserve Bank of New York.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The Role of Entrepreneurship in US Job Creation and Economic Dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2016). Where has all the skewness gone? the decline in high-growth (young) firms in the U.S. *European Economic Review*, 86:4–23.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2020). Changing Business Dynamism and Productivity: Shocks versus Responsiveness. *American Economic Review*, 110(12):3952–3990.

Dinopoulos, E. and Thompson, P. (1998). Schumpeterian Growth Without Scale Effects. *Journal of Economic Growth*, 3(4):313–335.

Doraszelski, U. and Jaumandreu, J. (2013). R&D and Productivity: Estimating Endogenous Productivity. *The Review of Economic Studies*, 80(4 (285)):1338–1383.

Dosi, G. (1988). Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature*, 26(3):1120–1171.

Dosi, G., Grazzi, M., Moschella, D., Pisano, G., and Tamagni, F. (2019). Long-term firm growth: An empirical analysis of US manufacturers 1959–2015. *Industrial and Corporate Change*, page dtz044.

Ericson, R. and Pakes, A. (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies*, 62(1):53–82.

Fernald, J. (2018). Is Slow Productivity and Output Growth in Advanced Economies the New Normal? *International Productivity Monitor*, (35):138–148.

Foster, L., Haltiwanger, J., and Krizan, C. (2001). Aggregate Productivity Growth: Lessons from Microeconomic Evidence. In *New Developments in Productivity Analysis*, Edited by C.R. Hulten, E.R. Dean, and M.J. Harper, pages 303–363. University of Chicago Press, Chicago.

Garcia-Macia, D., Hsieh, C.-T., and Klenow, P. J. (2019). How Destructive Is Innovation? *Econometrica*, 87(5):1507–1541.

Ghazi, S. (2019). Large Firms and Long-Run Growth. Manuscript.

Gutiérrez, G. and Philippon, T. (2019). The Failure of Free Entry. Working Paper 26001, National Bureau of Economic Research.

Ha, J. and Howitt, P. (2007). Accounting for Trends in Productivity and R&D: A Schumpeterian Critique of Semi-Endogenous Growth Theory. *Journal of Money, Credit and Banking*, 39(4):733–774.

Hall, B. H., Mairesse, J., and Mohnen, P. (2010). Chapter 24 - Measuring the Returns to R&D. In Hall, B. H. and Rosenberg, N., editors, *Handbook of the Economics of Innovation*, volume 2 of *Handbook of the Economics of Innovation, Volume 2*, pages 1033–1082. North-Holland.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2012). Who Creates Jobs? small versus Large versus Young. *The Review of Economics and Statistics*, 95(2):347–361.

Hopenhayn, H., Neira, J., and Singhania, R. (2018). From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share. Working Paper 25382, National Bureau of Economic Research.

Hopenhayn, H. A. (1992). Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica*, 60(5):1127–1150.

Hopenhayn, H. A. (2014). Firms, Misallocation, and Aggregate Productivity: A Review. *Annual Review of Economics*, 6(1):735–770.

Huergo, E. and Jaumandreu, J. (2004). Firms' age, process innovation and productivity growth. *International Journal of Industrial Organization*, 22(4):541–559.

Karahan, F., Pugsley, B., and Şahin, A. (2019). Demographic Origins of the Startup Deficit. Working Paper 25874, National Bureau of Economic Research.

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

Laincz, C. A. (2009). R&D subsidies in a model of growth with dynamic market structure. *Journal of Evolutionary Economics*, 19(5):643–673.

Laincz, C. A. and Peretto, P. F. (2006). Scale effects in endogenous growth theory: An error of aggregation not specification. *Journal of Economic Growth*, 11(3):263–288.

Lee, Y. and Mukoyama, T. (2015). Entry and exit of manufacturing plants over the business cycle. *European Economic Review*, 77:20–27.

Lotti, F., Santarelli, E., and Vivarelli, M. (2009). Defending Gibrat's Law as a long-run regularity. *Small Business Economics*, 32(1):31–44.

Madsen, J. B. (2008). Semi-endogenous versus Schumpeterian growth models: Testing the knowledge production function using international data. *Journal of Economic Growth*, 13(1):1–26.

Mand, M. (2019). On the cyclicality of R&D activities. *Journal of Macroeconomics*, 59:38–58.

Melitz, M. J. and Polanec, S. (2015). Dynamic Olley-Pakes productivity decomposition with entry and exit. *The RAND Journal of Economics*, 46(2):362–375.

Mukoyama, T. and Osotimehin, S. (2019). Barriers to Reallocation and Economic Growth: The Effects of Firing Costs.

Naude, W. (2020). From the Entrepreneurial to the Ossified Economy: Evidence, Explanations and a New Perspective.

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

Peretto, P. F. (2018). Robust endogenous growth. *European Economic Review*, 108:49–77.

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

Peretto, P. F. and Smulders, S. (2002). Technological distance, growth and scale effects. *The Economic Journal*, 112(481):603–624.

Restuccia, D. and Rogerson, R. (2017). The Causes and Costs of Misallocation. *Journal of Economic Perspectives*, 31(3):151–174.

Rossi-Hansberg, E. and Wright, M. L. J. (2007). Establishment Size Dynamics in the Aggregate Economy. *American Economic Review*, 97(5):1639–1666.

Sterk, V., Sedláček, P., and Pugsley, B. (2021). The Nature of Firm Growth. *American Economic Review*, 111(2):547–579.

Sutton, J. (1997). Gibrat's Legacy. *Journal of Economic Literature*, 35(1):40–59.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2):326–365.

Syverson, C. (2017). Challenges to Mismeasurement Explanations for the US Productivity Slowdown. *Journal of Economic Perspectives*, 31(2):165–186.

Thompson, P. (2001). The Microeconomics of an R&D-Based Model of Endogenous Growth. *Journal of Economic Growth*, 6(4):263–283.

Young, A. (1998). Growth without Scale Effects. *Journal of Political Economy*, 106(1):41–63.

# 6 Appendices

# A Stationary Model

I present the detrended version of the model, described by the following equations.

First, define $\forall X$, $\widetilde{X}_t = \frac{X_t}{Z_t^{\theta}}$; $\check{X}_t = \frac{X_t}{N_t^{\frac{1}{\epsilon-1}}}$; $\widehat{X}_t = \frac{X_t}{Z_t^{\theta} N_t^{\frac{1}{\epsilon-1}}}$.

The production function (9) is:

$$\widetilde{x}_{i,t} = z_{i,t}^{\theta} l_{x_{i,t}}^{\nu}, \tag{37}$$

where the first order condition for production labor is

$$l_{x_{i,t}} = \left[ \nu \frac{\epsilon - 1}{\epsilon \widehat{w}_t} \left( \frac{\widehat{y}_t}{n_t} \right)^{\frac{1}{\epsilon}} z_{i,t}^{\theta \frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon - \nu(\epsilon-1)}}, \tag{38}$$

where $y_t = \frac{Y_t}{L_t}$and for pricing:

$$\check{p}_{i,t} = \frac{\epsilon}{\epsilon - 1} \frac{\widehat{w}_t l_{x_{i,t}}^{1-\nu}}{\nu z_{i,t}^{\theta}}. \tag{39}$$

Plugging these into detrended dividend, it be re-expressed as a function of $z_{i,t}$ and $l_{Z_{i,t}}$ only:

$$\widehat{\pi}_{i,t}(z_{i,t}, l_{Z_{i,t}}) = \widehat{w}_t \left[ \frac{\epsilon}{(\epsilon - 1)\nu} - 1 \right] \left[ \nu \frac{\epsilon - 1}{\epsilon \widehat{w}_t} \left( \frac{\widehat{y}_t}{n_t} \right)^{\frac{1}{\epsilon}} z_{i,t}^{\theta \frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon - \nu(\epsilon-1)}} - \widehat{w}_t (l_{Z_{i,t}} + \Phi) \tag{40}$$

The stationary Bellman equation is:

$$\widehat{V}(z_{i,t}) = \max_{\{l_{Z_{i,t}}\}} \left\{ \widehat{\pi}_{i,t}(z_{i,t}, l_{Z_{i,t}}) + (1 + g_{t+1}^{Z})^{\theta} (1 + \lambda)^{\frac{1}{\epsilon-1}} \left( \frac{n_{t+1}}{n_t} \right)^{\frac{1}{\epsilon-1}} \frac{\widehat{c}_{t+1}}{\widehat{c}_t} \times \frac{1}{1 + r_{t+1}} \max\{\mathbb{E}_t \widehat{V}(z_{i,t+1}), 0\} \right\} \tag{41}$$

with the knowledge accumulation equation (12), whose stationary version is:

$$z_{i,t} = \frac{z_{i,t-1} + \alpha_{i,t} z_{i,t-1}^{\mu} l_{Z_{i,t-1}}^{\zeta}}{(1 + g_t^{Z})}, \tag{42}$$

The entry condition (19) is:

$$\mathbb{E}_t \widehat{v^E}_{i,t}(\alpha_{i,t+1}, z_{i,t+1}) \geq f_E, \tag{43}$$

The equilibrium conditions are modified as follows. The labor market clearing (23) becomes:

$$\frac{1}{n_t} = \frac{\int_0^{N_t} (l_{x_{i,t}} + l_{Z_{i,t}}) di}{N_t} + \Phi; \tag{44}$$

the law of motion of the number of establishment (20) is now:

$$\frac{n_{t+1}}{n_t} = \frac{1 + \frac{N_{E_t} - N_{X_t}}{N_t}}{1 + \lambda}; \tag{45}$$

output (22) is:

$$y_t = c_t + f_E \frac{N_{E_t}}{N_t} n_t; \tag{46}$$

## B Steady State Algorithm

Constuct a grid for the state $z$ (relative knowledge) and the shock $\alpha$ (productivity of R&D) by choosing respectively 190 and 85 grid points. The grid points are spaced in a way to obtain higher concentration for lower values, where non-linearities are present. Furthermore, construct another grid for $z$ and $\alpha$ used to capture more precisely the future expected states, with 673 and 505 equally spaced points between each consecutive point in the grids for the state and the shock.

Provide an initial guess for the detrended values of wage, output, number of firms and for the growth rate of average knowledge. These are the variables that firms take as given when making their decisions. I use a bisection method to update these guesses. Additionally, I provide an initial guess for the distribution of firms over the two firm-specific state variable and shock (relative knowledge and productivity of R&D).

Solve the firm's problem given by the detrended Bellman equation (41) via policy function iteration for the R&D labor of firms at each combination of grid points of the two state variables, subject to the constraint (42). As the choice is dependent on the expected value of $\alpha$ in the following period, I use Gauss-Hermite quadrature with 15 nodes to approximate the expectation, and I impose a non-negativity constraint for R&D labor. Furthermore, I compute the value of the firm after dividend payout. If this value is negative, R&D labor is set to 0, as the firm exits the market at the end of the period. To improve the precision of the expected values' estimates, I construct a new grid over and $\alpha$, with 6 equally spaced points between each consecutive point in the grids for the states, and a new grid for $z$ by choosing the grid points as the difference between the realized $z$ if $\alpha$ turns out to be the expected one, and each value of $\alpha$ represented in this new grid. I interpolate linearly the value function to obtain a matrix of values over these new grids for $z$ and $\alpha$.[11]

Solve for the expected value of entrants by using the value function computed above. As the value of entrants corresponds to the present value of next period firm value, the firm's decision depends on the expectation of the draw of $\alpha_{i,t+1}$ and $z_{i,t+1}$. This expectation is approximated by a Gauss-Hermitian quadrature with 15 nodes.

At this point, I find the beginning of the period stationary distribution given the guesses for the relevant aggregate variables. This is done by following these steps:

- From the previous period distribution, set the mass of firms at grid points for which firm value is negative to 0. I use the sum of the mass of remaining firms to compute the exit rate, before reweighting the distribution to ensure that the weights of continuing firms sum up to 1.
- Find the new distribution over $\alpha$, given the old distribution and the law of motion of $\alpha$.
- Find the new distribution of incumbents over $z_{i,t}$. This depends on the old distribution, R&D labor hired in the previous period at given $z_{i,t-1}$ and $\alpha_{i,t-1}$, on $z_{i,t-1}$, on $\alpha_{i,t-1}$ and on the probability of moving from the previous $\alpha$ to each $\alpha$ point on the grid.
- Compute the outer product of mass over $\alpha$ and $z$ to find the new distribution of incumbents over them.
- Find the distribution of firms that entered in the previous period over the state variable and shock, by drawing $z_{i,t}$ according to equation (17) and $\alpha_{i,t}$ according to equation (18).

[11]Importantly, as the function approximation needs to be good enough not only to interpolate, but also to extrapolate, since the large grid over $z$ includes much larger values than the initial $z$ grid, the highest and second highest values of alpha need to be high enough and chosen very close to each other in a way to ensure that the linear extrapolation is decently accurate.

- Find the entry rate as the sum of exit rate and population growth rate (the condition required to ensure stationarity in the number of firms, essentially imposing steady state) from equation (45).

- Compute the new mass of firms as the weighted average of the mass of incumbents and the mass of entrants, using the entry rate as the weight.

- Iterate until the mass of firms in every grid point is no further than 0.0001 from what it was in the previous iteration, or for a maximum of 4 (a number picked arbitrarilly to increase the speed of convergence) iterations.

Finally, the guesses of the aggregate variables need to be updated (I do so by using a bisection method and a step size of 0.0001). Find average production and R&D labor using the normalized distribution and the policy functions at each grid point. Compute the values output and number of firms from equations (46) and (44) respectively. Increase the wage if the left side of equation (43) is larger than the right side, and increase the growth rate of average knowledge if the distribution of establishments over $z$ is such that the average relative knowledge is larger than 1. Iterate until the values of consumption, number of firms, growth rate of average knowledge, wage, mass of firms over the state and shock and entry rate differ from the values obtained in the previous iteration by less than a tolerance level of 0.002, 0.001, 0.0006, 0.01, 0.0008, and 0.001 respectively.

# C Estimating the Gibrat Coefficient: Correlation Between Growth and Size

The model, in line with the data, violates what is known as Gibrat Law, a statement over the independence of the growth rate of size from its level. Specifically, a long tradition of empirical work, reviewed by Sutton (1997); Caves (1998); Audretsch et al. (2004), has devoted its focus on estimating the following equation:

$$\log SIZE_{t+1} - \log SIZE_t = \alpha + \beta \log SIZE_t + \epsilon_{t+1}, \tag{47}$$

where $\epsilon$ is the error term, and $\alpha$ and $\beta$ are parameters. Gibrat Law holds as long as $\beta = 0$. Size is typically measured in number of employees or sales in dollars.

I simulate the model and estimate equation (47) where size is captured as number of employees to find $\beta = -0.053$. Although the model rejects Gibrat Law, the coefficient is not very far from 0. This is broadly in line with the results obtained by Audretsch et al. (1999) on Italian manufacturing data, who reject Gibrat Law for most industries. In those industries where Gibrat Law is not rejected their confidence intervals include the model's the point estimate. Although the coefficient they obtain for some industries is further from zero than the one implied by the model, this is of little concern as my main goal is to avoid overestating the quantitative importance of scale dependence.

If I consider only the top 50% of firms in terms of size, and estimate Gibrat Law, I obtain $\beta = -0.025$, fairly close to the value $-0.005$ estimated by Bottazzi et al. (2007) for a sample medium and large firms in Italy.

# D Growth of Average Productivity, Business Dynamism, and the Firm Size Distribution

This section identifies the source of bias that arises out of disregarding scale dependence in productivity growth rates and the firm heterogeneity in their knowledge level. Furthermore, it presents a decomposition of aggregate productivity growth to illustrate two other contributions of the paper. First, disregarding scale dependence leads to misattributing part of growth as emerging from within firm growth rather than reallocation. Second, I show that the model allows for a structural interpretation of aggregate productivity growth decompositions common in the empirical literature. All components are interdependent as all variables are functions of the two state variables.

I conclude the section by presenting a quantitative account of the decomposition to illustrate the importance of churning in accounting for within-firm growth.

### D.1 Productivity Growth and Churning

Most existing endogenous growth models, after correcting for entry and exit, predict that the growth rate of average productivity equals the average growth rate across firms. This subsection illustrates that when productivity growth rates across firms are scale dependent, these models overestimate growth by the degree of mean-reversion.

This result is formalized in the following proposition, and it is proved in the rest of this subsection.

**Proposition 3** *If $\rho = 0$ , productivity growth rates across firms are scale dependent, and, under a time-invariant distribution of relative productivity, the growth rate of the average productivity of continuing firms is lower than the average of their growth rates by the degree of mean-reversion.*

I proceed in proving the proposition and in showing formally the interaction between growth of average productivity and churning. To simplify the exposition and to emphasize that this result does not depend on the weights attached to individual productivities, I focus on the unweighted average of relative productivity levels. To do that, I rely on the following lemma.

**Lemma 4** *Under a time-invariant distribution of relative productivity levels, the unweighted average of productivities grows at the same rate as the weighted average.*

This is proved by showing that weighted and unweighted average productivity have the same trend. Dividing the unweighted average of productivities by the weighted average yields $\frac{\overline{A}_t}{A_t} = \frac{1}{N_t}\int_0^{N_t} a_{i,t}di$, which is trendless when the relative productivity distribution is time-invariant.

The unweighted average of productivities, $\overline{A}_t$, can be decomposed in two alternative ways:

$$\overline{A}_t = \frac{1}{N_t}\int_0^{N_{c_{t-1}}} A_{i,t}di + \frac{1}{N_t}\int_0^{N_{E_{t-1}}} A_{i,t}di = \frac{1}{N_t}\int_0^{N_{c_t}} A_{i,t}di + \frac{1}{N_t}\int_0^{N_{X_t}} A_{i,t}di, \tag{48}$$

where $N_{c_{t-1}}$ is the number of firms active at time $t-1$ that continue operating at time $t$.

Taking growth rates yields:

$$g_t^{\overline{A}} = \underbrace{\frac{1}{N_t}\int_0^{N_{c_{t-1}}} g_t^{A_i}\,\overline{a}_{i,t-1}di}_{\text{Firm Productivity Growth}} + \underbrace{\left(\frac{N_{c_{t-1}}}{N_t} - \frac{N_{c_{t-1}}}{N_{t-1}}\right)\overline{a}_{c,t-1}}_{\text{Expanding Varieties}} + \underbrace{\left(1+g_t^{\overline{A}}\right)\frac{N_{c_{t-1}}}{N_t}\overline{a}_{E_t} - \frac{N_{X_{t-1}}}{N_{t-1}}\overline{a}_{X_t}}_{\text{Entry and Exit}}, \tag{49}$$

where $\bar{a}_{i,t} = A_{i,t}/\overline{A}_t$, $\bar{a}_{E_t} = \frac{1}{N_{E_{t-1}}} \int_0^{N_{E_{t-1}}} A_{i,t}/\overline{A}_t di$ and $\bar{a}_{X_t} = \frac{1}{N_{X_t}} \int_0^{N_{X_t}} A_{i,t}/\overline{A}_t di$. The growth of continuing firms can be further divided up to make explicit the role of the average growth rate and of the higher moments:

$$\frac{1}{N_t} \int_0^{N_{c_{t-1}}} g_t^{A_i} \, \bar{a}_{i,t-1} di = \frac{N_{c_{t-1}}}{N_t} \Bigg[ \underbrace{\bar{g}_t^{A_i}}_{\substack{\text{Average} \\ \text{Growth Rate}}} + \underbrace{\frac{1}{N_{c_{t-1}}} \int_0^{N_{c_{t-1}}} \left(a_{i,t} - \bar{a}_{c_t}\right) \left(g_t^{A_i} - \bar{g}_t^{A_i}\right)}_{\text{Mean Reversion}} \Bigg], \quad (50)$$

where $\bar{g}_t^{A_i}$ is the average growth rate, and the remaining factor is the covariance between relative (to the unweighted average) productivity level and growth rates. This covariance differs from 0 in the presence of scale dependence. Note that a covariance different from 0 implies the presence of churning in the economy, which arises if growth rates differ across firms. Nevertheless, this covariance is not accounting for all churning, but only for that churning that manifests itself in the form of mean-reversion. In fact, an economy characterized by churning, but without a mechanism that delivers scale-dependence would have a covariance between levels and growth rates of 0 and its long-run growth rate would be independent of the degree of churning.

Quantitatively, mean-reversion subtracts approximately 10% from the growth rate of the unweighted average productivity.

Now, consider an economy where growth rates are strictly decreasing in relative productivity conditional on R&D productivity as in this model with this calibration due to the mechanisms described above. Equations (33) and (34) show that R&D investment and growth rates differentials depend on the expected next period R&D productivity. If $\rho = 0$, the R&D productivity is *iid*, implying that investment and growth rates are independent of the current R&D productivity level, thus implying that the covariance between growth rates and relative productivity level equals the unconditional covariance between the two.

The implications of this propositions are worth discussing. The main takeaway is that focusing on understanding the determinants of the average growth rate across firms does not lead to an accurate understanding of the aggregate productivity growth process when growth rates exhibit scale-dependence. Starting from the same average growth rate across firms, models consistent with Gibrat Law or symmetric models would deliver a higher long-run growth rate than this model. Furthermore, notice that churning that arises out of mean-reversion in relative productivity is what generates a time-invariant non-degenerate distribution. Under calibrations that deliver a non-degenerate time-invariant firm size distribution, the interdependence between churning and average productivity growth is inevitable.

A more complex picture arises if we consider (as in this case) a situation where R&D productivity levels are positively autocorrelated. The synthetic data generated by the model may exhibit a correlation between relative productivity and R&D productivity. In this case, I rely on the Law of Total Covariance to illustrate the connection between the bias identified in this subsection, and the covariance between productivity growth and levels conditional on R&D productivity whose sign is restricted by the requirement that the distribution is non-degenerate. The covariance between productivity growth and levels can therefore be decomposed as follows:

$$cov\left(\mathbb{E}_{t-1}g_t^{A_i}, \overline{a}_{i,t-1}\right) = cov\left(\mathbb{E}_{t-1}g_t^{A_i}, \overline{a}_{i,t-1}|\alpha_{i,t-1}\right) + cov\left(\mathbb{E}\left[\mathbb{E}_{t-1}g_t^{A_i}|\alpha_{i,t-1}\right], \mathbb{E}\left[\overline{a}_{i,t-1}|\alpha_{i,t-1}\right]\right). \quad (51)$$

The first term on the right side of the equation is the degree of conditional mean reversion, which is negative if growth rates are strictly declining in the relative productivity level conditional on R&D productivity. The second term depends on the correlation between the relative productivity level and the R&D productivity. It is easy to see that if the two are uncorrelated, the second element in the covariance has no variation, implying a covariance of 0. If $\rho > 0$, this second term may be positive, thus countering the negative sign of the first term. This could happen if the correlation between $\alpha$ and $a$ is positive, but not strong enough to ensure that the negative effect on the growth rate of a higher relative productivity level is high enough to more than offset the positive effect of a higher R&D productivity. The process of entry and exit reduces this covariance between conditional growth rates and conditional relative productivity levels: exit removes firms with a low relative productivity level and a low ability to innovate, while entry introduces in the market firms that are not very productive initially, but have a higher than average ability to innovate. Given the calibration presented above, the correletion generated by the model between

$\alpha$ and $a$ is approximately 0, making the covariance between conditional growth rates and levels quantitatively neglegible.

### D.2 Growth of Average Productivity and Firm-Size Distribution

I now turn to showing that the economy's growth rate is interdependent to both the R&D productivity distribution and the distribution of relative technological knowledge due to in-house R&D and diminishing returns. The implication is that any parameter change that leads to a different distribution of either relative knowledge or R&D productivity, would necessarily deliver a different growth rate of average productivity. As a consequence, the two moments cannot be studied separately.

**Proposition 5** *If R&D is performed in house and there are diminishing returns per variety,* **A** *R&D investment is scale dependent and nonlinear in relative knowledge, and* **B** *R&D investment is nonlinear in the R&D productivity.*

The first part of the statement was shown in section 3, by noting that the policy function depends on the slope of the value function, which is decreasing in relative knowledge. The rest relies on the same explanation: due to diminishing returns per variety and in-house R&D, the firm value is concave in relative knowledge. Hence, the incentive to increase knowledge diminishes in itself, causing the marginal product of knowledge to be concave in the ability to innovate. The implication is that changes that deliver a different distribution of either relative knolwedge or R&D productivity would necessarily imply a different average productivity growth rate by changing incentives to innovate.

The growth rate of the unweighted average productivity (after a correction for entry, exit and change in the number of firms) is the sum of the productivity change of each firm detrended by the average productivity level, as shown in equation (49) in the previous subsection. This expression can be re-arranged to show what gives rise to the feedback mechanism from the relative knowledge (or productivity) and the R&D productivity distribution.

$$\frac{1}{N_t}\int_0^{N_{c_{t-1}}}\frac{A_{i,t}-A_{i,t-1}}{\overline{A}_{t-1}}di = \\ \frac{N_{t-1}}{\int_0^{N_{t-1}} z_{i,t-1}^{\theta}di}\frac{1}{N_{c_{t-1}}}\left[\int_0^{N_{c_{t-1}}}\left(z_{i,t-1}+\alpha_{i,t}z_{i,t-1}^{\mu}l_{Z_{i,t-1}}(.)^{\zeta}\right)^{\theta}-z_{i,t-1}^{\theta}di\right]. \tag{52}$$

As illustrated by this equation, the productivity change of each firm is a function of the product of three firm-specific terms. The first term is the exogenous R&D productivity shock. The second term is the firm-specific relative technological knowledge, which shows up due to the presence of private returns to own-knowledge in R&D. The third term is the R&D investment, which is itself a function of both the R&D productivity level (since the shock is persistent), and of relative knowledge.

How does this framework differ from one where R&D is independent of either R&D productivity, relative knowledge, or both? Answering this question would shed light on the role played by heterogeneity. I document the discrepancy between the growth of average productivity obtained in the partial equilibrium of this model, and various benchmarks where at least some of the channels through which heterogeneity manifests itself in this model are absent.

**Non-linear Scale Dependence in R&D**

As shown in section 3, R&D is a nonlinear function of relative knowledge, hence of productivity and firm-size. This nonlinearity would be of no interest in two types of models: a symmetric one where relative knowledge is the same for all firms, such as in Peretto and Connolly (2007); or a model where R&D is undirected, meaning that the firm is unaware of what product-line they improve over when investing, for example Acemoglu et al. (2018). If the firm-size dispersion increases, the partial equilibrium effect is that average R&D of firms with high innovative capacity decreases, while average R&D of firms with low innovative capacity increases. This is clear from figure 7 because around the average level of relative productivity, firms with a high innovative capacity are under the "bell-shape", while firms with a low innovative capacity are in the decreasing and convex part of the policy function. To quantify the extent of this discrepancy, I can compare the average R&D investment for firms with a given level of R&D productivity, and the R&D investment that would be observed under the average level of relative knowledge. This is visible in figure 6. Quantitatively, this channel alone leads to a wedge between the no-dispersion framework and this model by about 7% of the growth rate, with the latter higher than the former.

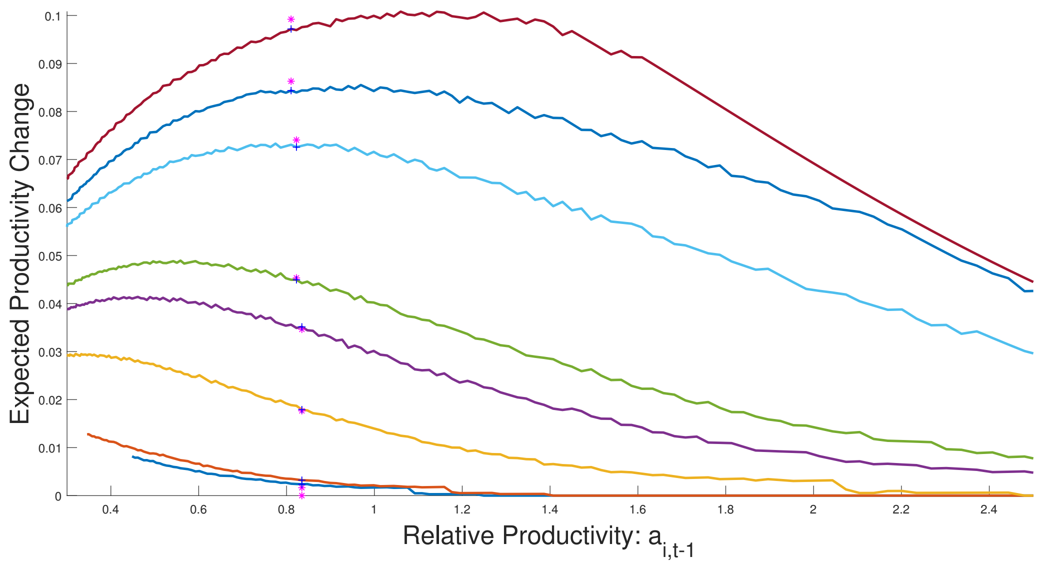


Figure 6: Expected productivity change for selected levels of R&D productivity. The stars represent the productivity change at the average level of R&D conditional on R&D productivity, while the pluses represent the average productivity change at the average productivity level.

### Non-linearity of R&D in R&D productivity

R&D is concave in R&D productivity due to diminishing returns per variety. This creates an interdependence between growth and the distribution of R&D productivity. This concavity is absent in models where there are constant returns per variety, such as those consistent with Gibrat Law, whose outcome is captured by the first moment of the R&D productivity distribution. As before, conditional on the level of relative knowledge, I compare the average R&D labor across firms with different R&D productivities to the R&D labor at the average R&D productivity. The concavity that arises ensures that a model based on the first moment of R&D productivity would have a 22% higher growth rate. This Jensen's inequality captures the discrepancy that arises from this channel between this model and models based on the first moment of R&D productivity in accounting for the long-run growth rate of the economy.

#### Diminishing returns to R&D in new knowledge production

Another obvious source of non-linearity arises from the presence of diminishing returns to R&D. This is a common feature to several growth models, such as Thompson (2001). Diminishing returns to R&D create an issue of aggregation of the firm specific R&D efforts such that the average R&D to the power $\zeta$ is lower than the average of the exponentiated individual R&D. This inequality arises whenever R&D effort differs across firms. Quantitatively, models where R&D is the same across firms would underestimate the average productivity's growth rate by 10%.

#### Diminishing Returns to own-knowledge in R&D

In this model, new knowledge creation depends linearly on overall knowledge. This is partly own-knowledge internal to the firm, and external knowledge. In symmetric models, own-knowledge and external knowledge are the same. Own knowledge in the R&D technology operates as if the R&D productivity was partly endogenous, emphasizing another channel through which the distribution of relative knowledge feeds back into the long-run growth rate of the economy.

#### Co-dependence of the Three Terms

The last element of interest is the co-dependence of the three terms. Considering all of them other than the returns to own-knowledge, a representative firm framework in partial equilibrium would deliver a growth of average productivity about 6 times higher than the one obtained with this model. If the returns to own-knowledge are considered as well, the discrepancy is even higher.

The main takeaway from this analysis is that one cannot discuss separately the firm-size distribution, the R&D productivity distribution, and the long-run growth rate of the economy. The first two affect the incentive to innovate, hence the contribution that each firm brings to the process of knowledge creation.

### D.3 Decomposing Average Productivity Growth

I show here a decomposition of average productivity growth decomposition based on those common in the growth accounting literature (Foster et al., 2001; Melitz and Polanec, 2015). I explicitly show the role of the higher order moments when firms' growth rates exhibit negative scale dependence. In doing so, I point to the misattribution bias implied by models where growth rates are independent of the relative productivity level. Furthermore, I elucidate on the interdependence of the various components.

Aggregate productivity growth is decomposed into an accounting identity using the Foster et al. (2001) decomposition. The dynamic Olley Pakes decomposition proposed by Melitz and Polanec (2015) is instead discussed in the following section. The change in aggregate productivity is:

$$
\begin{aligned}
A_t - A_{t-1} = & \underbrace{\int_0^{N_{c,t-1}} S_{i,t-1}\Big(A_{i,t} - A_{i,t-1}\Big) di}_{\text{Within}} + \underbrace{\int_0^{N_{c,t-1}} \Big(S_{i,t} - S_{i,t-1}\Big)\Big(A_{i,t-1} - A_{t-1}\Big) di}_{\text{Between}} \\
& + \underbrace{\int_0^{N_{c,t-1}} \Big(S_{i,t} - S_{i,t-1}\Big)\Big(A_{i,t} - A_{i,t-1}\Big) di}_{\text{Cross-Correlation}} \\
& + \underbrace{\int_0^{N_{E,t-1}} S_{i,t}\Big(A_{i,t} - A_{t-1}\Big) di}_{\text{Entry}} - \underbrace{\int_0^{N_{X,t-1}} S_{i,t-1}\Big(A_{i,t-1} - A_{t-1}\Big) di}_{\text{Exit}}, \quad (53)
\end{aligned}
$$

where $N_{c,t-1}$ denotes the number of firms operating at time $t-1$ that continue their operations at time $t$, $N_{X,t-1}$ is the number of firms that exit at time $t-1$, and $N_{E,t-1}$ is the number of firms that enter at time $t-1$ and starts operating at time $t$.

Equation (53) can be re-expressed in proportional growth rates by dividing both sides by $A_{t-1}$. This leads to:

$$
\begin{aligned}
g_t^A = & \underbrace{\int_0^{N_{c_{t-1}}} S_{i,t-1}\ a_{i,t-1}\ g_t^{A_i} di}_{\text{Within}} + \underbrace{\int_0^{N_{c_{t-1}}} \Big(S_{i,t} - S_{i,t-1}\Big)\Big(a_{i,t-1} - 1\Big) di}_{\text{Between}} \\
& + \underbrace{\int_0^{N_{c_{t-1}}} \Big(S_{i,t} - S_{i,t-1}\Big) a_{i,t-1}\ g_t^{A_i} di}_{\text{Cross-correlation}} \\
& + \underbrace{\int_0^{N_{E_{t-1}}} S_{i,t}\Big(\frac{A_t}{A_{t-1}} a_{i,t} - 1\Big) di}_{\text{Entry}} - \underbrace{\int_0^{N_{X_{t-1}}} S_{i,t-1}\Big(a_{i,t-1} - 1\Big) di}_{\text{Exit}}
\end{aligned}
\quad (54)
$$

where $1 + g_t^{A_i} = (1 + g_t^{Z_i})^\theta$.

This decomposition highlights the contribution of different sources to aggregate productivity growth. The first component is the within-firm growth rate; it keeps track of the change in productivity of each surviving firm. A detailed description is provided below. The between component describes the contribution to growth of a reallocation of market shares between firms. Its sign is positive if business stealing favors on average firms with a higher initial relative productivity, and vice-versa. The cross-correlation component is a correction term that captures the correlation between productivity change and change in market share. It accounts for the gain, in terms of market share, of an increase in productivity. Instead, the last two components are indicators of the static importance of entry and exit respectively to the aggregate growth process.

The accounting identity presented in equation (68) is not new. One of this paper's contributions is to offer a theoretical interpretation of it, where all components are endogenously and jointly determined. In fact, I can express them as functions of the same variables, namely relative knowledge and firm-specific productivity of R&D.

$$a_{i,t} = a_{i,t}(z_{i,t}). \tag{55}$$

Firm-specific productivity is increasing in its knowledge stock by definition. Relative productivity is a function of relative knowledge.

$$N_t = N_t(z_t, \alpha_t), \tag{56}$$

where $z_t$ is a vector of all firms' relative knowledge and $\alpha_t$ a vector of the R&D productivities. The number of firms is endogenously determined and depends on each firm's entry and exit decision, hence on their state.

$$S_{i,t} = S_{i,t}(a_{i,t}(z_{i,t}), N_t(z_t, \alpha_t)). \tag{57}$$

The labor share increases in each firm's relative productivity level and decreases in the number of firms in the market.

$$A_t = A_t(Z_t, S_t(.)). \tag{58}$$

The average productivity level is increasing in the knowledge level of each firm and depends on their labor share.

$$g_t^{A_i} = g_t^{A_i}(z_{i,t-1}, \alpha_{i,t}, l_{Z_{i,t-1}}(z_{i,t-1}, \alpha_{i,t-1}, a_{i,t}(.), S_{i,t}(.))). \tag{59}$$

Finally, as discussed above, the firm-specific growth rate of productivity depends positively on the realization of the shock $\alpha_{i,t}$, on the existing stock of knowledge as private knowledge enhances the ability to produce new knowledge, and on their R&D investment. Nevertheless, R&D is itself a function of both relative knowledge and R&D productivity. It increases with the current R&D productivity, while the dependence on relative knowledge is instead more intricated. On the one hand, the higher the private stock of knowledge, the higher the ability to produce new ideas, hence the higher the incentive to innovate. On the other hand, the higher the knowledge stock, the higher the productivity level and the labor share. Consequently, due to the presence of diminishing returns per variety, the return to investment decreases as firms increase their relative productivity level and their market share, holding all else constant.

As all variables are endogenously and jointly determined, the decomposition cannot be interpreted as if market share, productivity level, and growth were random variables. It is not possible to ask what effect a change in one of the components has on aggregate growth while holding everything else constant because changing one variable would necessarily change others. Changes in the various components have common causes.

In what follows, I analyze further each component, imposing theoretical restrictions on them.

**Within Component**

Here I decompose the within component further to show explicitly the contribution of the higher moments of the distribution to within-firm growth.

One can notice from (54) that the within component is the average of the growth rates across firms weighted twice, first by labor share and then by relative productivity. The labor share is used as a weight by definition. Growth rates are further weighted by relative productivity as shown in the previous section due to an accounting property of arithmetic average.

Define the combined labor share of continuing firms as $S_{c_t} = \int_0^{N_{c_t}} S_{i,t} di$ and their average productivity level as $A_{c_t} = \int_0^{N_{c_{t-1}}} \frac{S_{i,t}}{S_{c_t}} A_{i,t} di$. It follows that:

$$\int_0^{N_{c_{t-1}}} \frac{1}{N_{c_{t-1}}} \left( S_{i,t} a_{i,t-1} - S_{c_t} a_{c_{t-1}} \right) di = 0. \tag{60}$$

Given this, the within component can be divided further:

$$\begin{aligned} g_t^{WITHIN} = & \underbrace{\overline{g}_t^{A_i}(.)}_{\text{Average growth rate}} + \underbrace{\left(S_{c_{t-1}}(.) a_{c_{t-1}}(.) - 1\right) \overline{g}_t^{A_i}(.)}_{\text{Survival selection}} \\ & + \underbrace{\int_0^{N_{c_{t-1}}(.)} \left( S_{i,t-1}(.) a_{i,t-1}(.) - \frac{S_{c_t}(.)}{N_{c_{t-1}}(.)} a_{c_{t-1}}(.) \right) \left( g_t^{A_i}(.) - \overline{g}_t^{A_i}(.) \right)}_{\text{Systematic Churning}}, \end{aligned} \tag{61}$$

where $\overline{g}_t^{A_i} = \int_0^{N_{c_{t-1}}} \frac{1}{N_{c_{t-1}}} g_t^{A_i} di$ represents the simple average of the productivity growth rates of individual continuing firms.

The *survival selection* component is a correction term. It has to do with the relative productivity level of surviving firms. Its sign and size depend on whether surviving firms are on average more, less, or equally productive to those that exit. The higher the relative productivity of survivors, the higher the contribution to aggregate growth from this channel. Models that assume an exogenous exit shock that hits all firms symmetrically would miss this.

The last component captures the degree of mean reversion, which indicates the presence of what I refer to as *systematic churning*. It differs from zero if growth rates are correlated to firm size. In fact, this component is a covariance between growth rates and the product of labor share and relative productivity. As labor share is a strictly increasing function of relative productivity, systematic churning shows up as a covariance between a monotonic transformation of relative productivity and the productivity growth rate. The word *systematic* is meant to stress that this element differs from zero only if growth rates differ in expectation. As the realization of the shock is independent of the relative productivity level, this component can differ from zero only in the case of investment decisions that lead to different predicted growth rates depending on the initial size, namely if firms find it optimal to move to a different position within the relative productivity distribution.

**Between Component**

I now turn to showing what makes up the between component, which tracks the contribution of a reallocation of labor share to firms with a higher productivity level.

$$g_t^{BETWEEN} = \underbrace{\int_0^{N_{c_{t-1}}(.)} \left(\Delta S_{i,t}(.) - \frac{\Delta S_{c_t}(.)}{N_{c_{t-1}}(.)}\right)\left(a_{i_{t-1}}(.) - \bar{a}_{c_{t-1}}(.)\right) di}_{\text{Static Reallocation}} + \underbrace{\Delta S_{c_t}(.)\left(\bar{a}_{c_{t-1}}(.) - 1\right)}_{\text{Reallocation to Incumbents}}. \quad (62)$$

The last term is a correction for the change in the combined labor share of incumbents, driven by the presence of entry and exit. The more interesting term is the first one, which is an additional manifestation of the main theme of this paper. It shows the covariance between the change in labor share and the initial relative productivity level. As the labor share is an increasing function of the relative productivity level, the negative scale dependence discussed so far implies a negative sign for this covariance. The between component, after a correction for the effect of entry and exit, is therefore negative in steady state.

**Cross-Correlation Component**

Here, I show a decomposition of the cross-correlation component. While the between component shows the gains for average productivity of a reallocation of labor to more productive firms, the cross-correlation component illustrates the contribution of a reallocation of labor to firms whose productivity level grows at a faster rate.

The cross-correlation component is therefore broken down into:

$$g_t^{CC} = \underbrace{\int_0^{N_{c_{t-1}}(.)} \left(\Delta S_{i,t}(.) - \frac{\Delta S_{c_t}(.)}{N_{c_{t-1}}(.)}\right) \left(1 + g_t^{A_i}\right) a_{i_{t-1}}(.)di}_{\text{Dynamic Reallocation}}$$
$$- \underbrace{\int_0^{N_{c_{t-1}}(.)} \left(\Delta S_{i,t}(.) - \frac{\Delta S_{c_t}(.)}{N_{c_{t-1}}(.)}\right) \left(a_{i_{t-1}}(.) - \bar{a}_{c_{t-1}}(.)\right) di}_{\text{Static Reallocation}}$$
$$+ \underbrace{\Delta S_{c_t}(.) \frac{1}{N_{c_{t-1}}(.)} \int_0^{N_{c_{t-1}}} \left(\frac{A_{i,t} - A_{i,t-1}}{A_{t-1}}\right) di}_{\text{Reallocation to Incumbents}} . \quad (63)$$

The second element is the same that appears in the between component, but with the opposite sign, implying that in steady state static reallocation adds to the cross-correlation component. The third factor depends on whether the labor share of continuing firms expands or not. The first one is instead the interesting one. It shows the contribution to average productivity growth of the reallocation of labor share to firms that grow the most. It can be thought of as a covariance between the change in market share and the productivity change detrended by average productivity. It is useful to further divide it up as follows:

$$\int_0^{N_{c_{t-1}}} \left(\Delta S_{i,t-1}(.) - \frac{\Delta S_{c_t}(.)}{N_{c_{t-1}}}\right) \left(g_t^{A_i} - g_t^{\bar{A}_i}\right) a_{i_{t-1}}(.)di =$$
$$\underbrace{\int_0^{N_{c_{t-1}}} \left(\Delta S_{i,t}(.) - \frac{\Delta S_{c_t}(.)}{N_{c_{t-1}}}\right) \left(\mathbb{E}_{t-1} g_t^{A_i} - g_t^{\bar{A}_i}\right) a_{i_{t-1}}(.)di}_{\text{Systematic Dynamic Reallocation}}$$
$$+ \underbrace{\int_0^{N_{c_{t-1}}} \left(\Delta S_{i,t}(.) - \frac{\Delta S_{c_t}(.)}{N_{c_{t-1}}}\right) \left(g_t^{A_i} - \mathbb{E}_{t-1} g_t^{A_i}\right) a_{i_{t-1}}(.)di}_{\text{Unsystematic Dynamic Reallocation}} \quad (64)$$

Part of the dynamic reallocation is therefore the result of the systematic churning process, while the rest is due to the realization of the shock, which leads firms that receive a positive draw to expand their labor share at the expense of those that get a negative one. Additionally, one can separate the systematic dynamic reallocation as the covariance between *today*'s labor share and detrended productivity change, minus the covariance between *yesterday*'s labor share and detrended productivity change. As it turns out, this last element resambles systematic churning as it appears in the within component with the opposite sign.[12] The absence of scale dependence would eliminate these terms and is therefore responsible for a misattribution bias, in favor of the within-firm component and against the cross-correlation component.

**The Role of Entry and Exit**

Entry and exit play a role in shaping firms' distribution over relative knowledge and R&D productivity, affecting the average growth rate, churning, and the aggregate growth process.

A selection effect driven by entry and exit affects the cross-sectional average of R&D productivity. As entrants have a higher ability to innovate than exiting firms on average, the cross-sectional average is higher than the long-run expectation of the AR(1) process. This selection effect is central in the model presented by Acemoglu et al. (2018) and has consequences for within-firm growth.

Another important selection effect concerns the relative productivity level, as emphasized by the firm dynamics literature (Syverson, 2011). Both entering and exiting firms have a below-average productivity level, but in steady state more exit also implies more entry. Therefore, the net direct effect on the aggregate productivity level caused by entry and exit is due to the difference in productivity levels between entrants and exiting firms, weighted by market share.

I now explore what concerns the dispersion of productivity and growth rates, which has implications for aggregate growth in this model. In particular, as systematic churning is a covariance between a monotonic transformation of labor share and productivity growth rates, entry can reinforce that if it increases the dispersion of growth rates, of relative productivity levels, or the correlation between these two variables.

Since firms that exit have a low level of both productivity and R&D productivity, exit reduces the dispersion in both state variables. This is clearly visible from figure 3.

The effect of entry is instead less obvious. Entrants have a higher dispersion of R&D productivity. The share of entrants that exit is higher than for all firms and the share of firms that grow faster than average is also higher. To better understand the effect of entry on churning, I need to focus on those entrants who decide to stay in the market for one more year. The average relative productivity for continuing entrants is 0.78, their standard deviation is 0.11, compared to 0.23 for all firms.

[12] The two are not exactly equal to each other: their difference is indeed captured by the covariance between productivity growth and level discussed in the previous section.

This points to an ambiguous change in the dispersion of relative productivity. On the one hand, as the average productivity level is far from that of incumbents, entry increases the productivity dispersion. On the other, the low dispersion of relative productivity compared to incumbents reduces the aggregate level. Comparing firms' productivity dispersion with and without entrants, I obtain the same number, implying that entry does not directly change the dispersion.

The effect on growth rates is instead different. The average productivity growth rates across incumbents, 3%, and across entrants, 5%, are both targeted moments. Their standard deviations are 0.14 and 0.10 respectively. Again, the dispersion of growth rates is ambiguously affected by entry. The standard deviation of growth rates when disregarding entrants is 0.08.

Besides, as entry introduces firms with below-average relative productivity and above-average growth rates, the correlation between the two is likely to increase in absolute value. I can compare the dispersion of the level and growth rate of productivity and their correlation for all firms while including and excluding entrants to find $-0.10$ and $-0.11$ respectively. The absolute value of the correlation is increased by entry, but the effect appears to be small. Furthermore, as already pointed out earlier, both entry and exit reduce the correlation between the relative productivity level and the R&D productivity. This has a direct negative effect on the covariance between relative productivity level and productivity growth as mentioned in the discussion over equation (51).

These results show that one dimension that affects the magnitude of systematic churning is increased by entry, while the other two are left almost unaltered. In conclusion, given this calibration, entry increases systematic churning. Additionally, firms that tend to grow faster are indeed young. The negative correlation that drives systematic churning and violates Gibrat Law depends crucially on age, in line with empirical evidence (Lotti et al., 2009; Haltiwanger et al., 2012).

### D.4 Quantitative Analysis of the Growth Decomposition

As common in the literature, for example, in Acemoglu et al. (2018), I present a quantitative analysis of the aggregate growth decomposition based on data generated by the model.

Table 8 shows the decomposition results presented in equation (54), where all components are shown as shares of average productivity growth.

Table 8: Productivity growth decomposition as a fraction of the growth rate of average productivity.

| Within | Between | Cross-Correlation | Net Entry |
|---|---|---|---|
| 42% | -27% | 78% | 6% |

Table 9 presents the contribution, as a share of the within component, of the elements that determine the change in surviving firms' average productivity. This is based on equation (61). As the results show, the effect of systematic churning is quantitatively relevant. Overall, the average of the growth rates is over 70% higher than the average productivity growth rate. This suggests that, in steady state, heterogeneity in firms' growth rates is important, and the bias attributable to disregarding it can be quite big. Systematic churning subtracts almost 70% from the within component and 40% from aggregate productivity growth.

Table 9: Within Component Decomposition. All elements are expressed as a share of the growth rate of average productivity.

| Average Growth | Systematic Churning | Survival Selection |
|---|---|---|
| 77% | -32% | -3% |

Another interesting result is that one-year-old firms account for a disproportionate share of systematic churning, around 14% of it, while representing only 6% of the total number of surviving plants, weighted by market share. This accounts for a share of the within component of $-9\%$, while their average growth rate accounts for 24% of it. This implies that overall one-year-old firms account for 14% of the within growth rate. In sum, they affect disproportionately the within-firm growth rate, but less than an observer would perceive by focusing exclusively on their average growth rate.

Table 10: Cross-Correlation Component Decomposition. All elements are expressed as a share of the growth rate of average productivity. Note that the static reallocation component enters negatively in cross-correlation.

| Static Reallocation | Dynamic Reallocation | Reallocation to Incumbents |
|---|---|---|
| -38% | 42% | -2% |

The dynamic reallocation component is by far the largest one. Due to the absence of any friction to the reallocation of labor, this is likely to be overestimated. The systematic churning component is quite important too, even though the estimate of Gibrat Law is quite conservative and the amount of reallocation is underestimated.

# E Disentangling Research Productivity

This section is devoted to exploring the determinants of the economy's ability to innovate. The average growth component can be further dissected to explore the connection between firms' average growth rate and average R&D labor:

$$
\begin{aligned}
&\bar{g}_t^{A_i} = \\
&\underbrace{\frac{1}{N_{c_{t-1}}}\left[\int_0^{N_{c_{t-1}}} \left(1+\alpha_{i,t} z_{i,t-1}^{\mu-1} l_{Z_{i,t-1}}(z_{i,t-1},\alpha_{i,t-1})^{\zeta}\right)^{\theta} - \left(1+\alpha_{i,t} l_{Z_{i,t-1}}(z_{i,t-1},\alpha_{i,t-1})^{\zeta}\right)^{\theta} di\right]}_{\text{Knowledge diffusion}} \\
&+ \underbrace{\frac{1}{N_{c_{t-1}}}\left[\int_0^{N_{c_{t-1}}} \left(1+\alpha_{i,t} l_{Z_{i,t-1}}(z_{i,t-1},\alpha_{i,t-1})^{\zeta}\right)^{\theta} - \left(1+\alpha_{i,t} l_{Z_{i,t-1}}(z_{i,t-1},\alpha_{i,t-1})\right)^{\theta} di\right]}_{\text{R\&D Aggregation}} \\
&+ \underbrace{\int_0^{N_{c_{t-1}}} \frac{\left(1+\alpha_{i,t} l_{Z_{i,t-1}}(z_{i,t-1},\alpha_{i,t-1})\right)^{\theta}}{N_{c_{t-1}}} - \left[1+\int_0^{N_{c_{t-1}}} \frac{\alpha_{j,t}}{N_{c_{t-1}}} dj \int_0^{N_{c_{t-1}}} \frac{l_{Z_{j,t-1}}(z_{j,t-1},\alpha_{i,t-1})}{N_{c_{t-1}}} dj\right]^{\theta} di}_{\text{R\&D Allocation}} \\
&+ \left[1+ \underbrace{\int_0^{N_{c_{t-1}}} \frac{\alpha_{i,t} di}{N_{c_{t-1}}}}_{\text{Average R\&D Productivity}} \underbrace{\int_0^{N_{c_{t-1}}} \frac{l_{Z_{i,t-1}}(z_{i,t-1},\alpha_{i,t-1})}{N_{c_{t-1}}} di}_{\text{Average R\&D}}\right]^{\theta} - 1. \qquad (65)
\end{aligned}
$$

Equation (65) shows that the average productivity growth rate is not exclusively explained by the average R&D effort and the average R&D productivity. Instead, it depends on a few other elements related to the higher moments of the distribution of relative knowledge and firm-specific research productivity. As a result, attempts to capture the economy's ability to innovate by computing a ratio of the average growth rate and average R&D include biases that mask the heterogeneity in firms' growth process.

The first component keeps track of the contribution of the knowledge spillover to productivity growth. Depending on the size of the parameter $\mu$, which describes the returns to private knowledge in the process of knowledge accumulation, firms with a relatively lower stock of knowledge will be growing faster. The knowledge spillover operates as a force of attraction to the average knowledge stock.

The second component highlights an issue of aggregation of R&D labor that arises due to diminishing returns to R&D in the knowledge production function. This component could be further divided up into two parts: one keeping track of the bias arising from disregarding the concavity and the other capturing the aggregation bias associated with the failure to account for R&D's dispersion across firms.

The last component is a correction of first-moment models' failure to associate R&D effort to the firm that performs it. Its magnitude depends on the dispersion of the ability to innovate across firms and its association with R&D effort.

The next subsection provides the quantitative results from equation (65).

## E.1 Research Productivity Decomposition

I present the quantitative results associated with the decomposition presented in the previous section. These results highlight the factors that matter in accounting for the average ability to innovate across firms.

Table 11 shows the decomposition of the average growth rate, according to equation (65). The various components are expressed as shares of the average growth rate.

Table 11: Research Productivity Decomposition

| First Moments | Knowledge diffusion | Aggregation | Allocation |
|---|---|---|---|
| 587% | 59% | -10% | -536% |

The results show that the first moments alone, namely the average of R&D labor and the average of firms' research productivity, provide a very poor estimate of the average growth rate. The average growth rate is overestimated by almost 6 times. The component that matters the most is the allocation of R&D effort. As the research productivity is widely dispersed, it is important to take into account where the research effort is directed to get an accurate description of the economy's ability to promote technological change. This suggests that phenomena that may affect the allocation of R&D, such as financing constraints, are of great importance and should be investigated further.

Another relevant issue to keep in mind is that the catching up component, which has to do with the knowledge spillover effect, explains a substantial portion of firm growth. The reason why its sign is positive is that most firms are below the average knowledge level, given the skewness of the distribution. Note that entrants affect disproportionately this component, accounting for 19% of it. This result highlights once again the relevance of life cycle dynamics for understanding the drivers of aggregate growth, and it shows that entrants grow faster than incumbents partly due to a spillover effect.

## F Dynamic Olley-Pakes Decomposition

Here, I show the decomposition proposed in Melitz and Polanec (2015). Although less common than the one presented above, it may nevertheless prove useful as the within component is independent of market shares, which allows me to illustrate more clearly the contrast with the theoretical growth literature. The change in aggregate productivity is:

$$A_t - A_{t-1} = \int_0^{N_{c_{t-1}}} \frac{A_{i,t}}{N_{c_{t-1}}} di - \int_0^{N_{c_{t-1}}} \frac{A_{i,t-1}}{N_{c_{t-1}}} di + \Delta cov_{c_t} + S_{E_t}\left(A_{E_t} - A_{c_t}\right) + S_{X_t}\left(A_{c_{t-1}} - A_{X_{t-1}}\right), \quad (66)$$

where

$$\Delta cov_{c_t} = \int_0^{N_{c_{t-1}}} \left(\frac{S_{i,t}}{S_{c_t}} - \frac{1}{N_{c_{t-1}}}\right)\left(A_{i,t} - A_{c_t}\right) di - \int_0^{N_{c_{t-1}}} \left(\frac{S_{i,t-1}}{S_{c_{t-1}}} - \frac{1}{N_{c_{t-1}}}\right)\left(A_{i,t-1} - A_{c_{t-1}}\right) di \quad (67)$$

Equation (66) can be re-expressed in proportional growth rates by dividing both sides by $A_{t-1}$. The dynamic Olley Pakes decomposition is therefore:

$$g_t^A = \underbrace{\int_0^{N_{c_{t-1}}} \frac{1}{N_{c_{t-1}}} a_{i,t-1}\, g_t^{A_i}}_{\text{Within}} + \underbrace{\frac{\Delta cov_{c_t}}{A_t}}_{\text{Between}} + \underbrace{S_{E_t} \frac{A_t}{A_{t-1}}\left(a_{E_t} - a_{c_t}\right)}_{\text{Entry}} + \underbrace{S_{X_{t-1}}\left(a_{c_{t-1}} - a_{X_{t-1}}\right)}_{\text{Exit}}, \quad (68)$$

where $a_{i,t} = \frac{A_{i,t}}{A_t}$, $a_{c_t} = \frac{A_{c_t}}{A_t}$, $a_{E_t} = \frac{A_{E_t}}{A_t}$, $a_{X_t} = \frac{A_{X_t}}{A_t}$, and $1 + g_t^{A_i} = (1 + g_t^{Z_i})^{\theta}$.

This decomposition, highlights the contribution of different sources to aggregate productivity growth. The first component is the within-firm growth rate; it keeps track of the change in absolute productivity levels of all continuing firms. The second one describes the contribution to growth of a reallocation of market shares between firms. The third and fourth are instead indicators of the static importance of entry and exit respectively to the aggregate growth process, and they could be further divided up into a within and a between component themselves.

I now show where the same two contributions presented above show up in this decomposition. First, I can decompose the within component further by dividing it up in two parts, an average growth and a heterogeneity component. Second, I can express relative productivity, growth rates, market shares and number of surviving, entering or exiting establishments as functions of the same variables, namely relative knowledge and the firm-specific productivity of R&D. With these modifications, equation (68) becomes:

$$
\begin{aligned}
g_t^A(z_{t-1}, z_t, \alpha_t, \mathbb{E}_{t-1}\alpha_t) = \\
&\underbrace{\underbrace{\overline{g}_t^{A_i}(z_{t-1}, \alpha_t, \mathbb{E}_{t-1}\alpha_t)}_{\text{Average growth}} + \underbrace{\int_0^{N_{c_{t-1}}(z_{t-1})} \frac{a_{i,t-1}(z_{i,t-1}) - 1}{N_{c_{t-1}}(z_{t-1})} g_t^{A_i}(z_{i,t-1}, \alpha_{i,t}, \mathbb{E}_{t-1}\alpha_{i,t}) di}_{\text{Heterogeneity}}}_{\text{Within}} \\
&+ \underbrace{\frac{\Delta cov_{c_t}(z_t, z_{t-1})}{A_t(z_t)}}_{\text{Between}} - \underbrace{\frac{cov\big(A_{i,t-1}(z_{i,t-1}), S_{i,t-1}(z_{i,t-1})\big)}{\overline{A}_{t-1}(z_{t-1})} g_t^{WITHIN}}_{\text{Productivity Reward (cross-correlation)}} \\
&\underbrace{+ S_{E_t}(z_t) \frac{A_t(z_t)}{A_{t-1}(z_{t-1})} \Big(a_{E_t}(z_t) - a_{c_t}(z_t)\Big)}_{\text{Entry}} + \underbrace{S_{X_{t-1}}(z_{t-1}) \Big(a_{c_{t-1}}(z_{t-1}) - a_{X_{t-1}}(z_{t-1})\Big)}_{\text{Exit}}, \qquad (69)
\end{aligned}
$$

where $z_t$ is a vector of the relative knowledge of all firms and $\alpha_t$ a vector of the research productivities, and $\overline{g}_t^{A_i} = \int_0^{N_{c_{t-1}}} \frac{1}{N_{c_{t-1}}} g_t^{A_i} di$ represents the simple average of the productivity growth rates of individual continuing establishments.

### Understanding Heterogeneity

Heterogeneity enters the picture in the form of a covariance and of a correction term that depends on selection among exiting firms. In what follows, I present a decomposition of the heterogeneity component to introduce the moments that are relevant in accounting for aggregate growth.

Using the fact that $\int_0^{N_{c_{t-1}}} \frac{1}{N_{c_{t-1}}} \Big(\overline{a}_{i,t-1} - \overline{a}_{c_{t-1}}\Big) di = 0$, where $\overline{a}_{i,t} = \frac{A_{i,t}}{\overline{A}_t}$ and $\overline{a}_{c_t} = \frac{\overline{A}_{c_t}}{\overline{A}_t}$, and $\overline{A}_{i,t} = \int_0^{N_t} \frac{A_{i,t}}{N_t} di$, the heterogeneity component can be divided further to better capture all elements that it includes:

$$g_t^{HETEROGENEITY} = + \underbrace{\Big(\overline{a}_{c_{t-1}}(z_{i,t-1}) - 1\Big)\overline{g}_t^{A_i}(z_t, \alpha_t, \mathbb{E}_{t-1}\alpha_t)}_{\text{Survival selection}} + \underbrace{cov\big(g_t^{A_i}(z_{i,t}, \alpha_{i,t}, \mathbb{E}_{t-1}\alpha_{i,t}), \overline{a}_{i,t-1}(z_{i,t-1})\big)}_{\text{Systematic churning}}. \quad (70)$$

Here, systematic churning is a covariance between relative productivity level and absolute productivity growth rates across firms:

$$cov(g_t^{A_i}(.), \overline{a}_{i,t-1}(z_{i,t-1})) = \int_0^{N_{c_{t-1}}(z_{t-1})} \frac{1}{N_{c_{t-1}}(z_{t-1})} \left(g_t^{A_i}(.) - \overline{g}_t^{A_i}(.)\right) \Big(\overline{a}_{i,t-1}(z_{i,t-1}) - \overline{a}_{c_{t-1}}(z_{t-1})\Big) di. \quad (71)$$

It describes the nature of firm mobility within the relative productivity distribution. This component is different from zero in the presence of heterogeneity in productivity growth rates across firms, in productivity levels, and in the correlation between the two.

Equation (71) shows that churning does indeed create a wedge between the growth rate of the within component and the average growth rate across firms even when the within component is defined as an unweighted mean of individual productivities. Nevertheless, defining aggregate productivity as a geometric mean, as opposed to an arithmetic mean, would lead to the disappearance of heterogeneity from the decomposition, thus hiding the contribution of churning.

**Quantitative Results**

Table 12 shows the results from the decomposition presented in equation (69), where all components are shown as shares of aggregate productivity growth, and equation (70), where all components enter as shares of the within component.

Table 12: Aggregate Growth Decomposition. All elements are expressed as a share of the growth rate of aggregate productivity.

| Within | Cross-Corr | Between | Net Entry |
|---|---|---|---|
| 92% | -11% | 23% | -4% |

Table 13 presents the contribution, as a share of the within component, of the elements that determine the change in average productivity of surviving firms.

Table 13: Within Component Decomposition. All elements are expressed as a share of the within component.

| Average Growth | Systematic Churning | Survival Selection |
|---|---|---|
| 114% | -15% | 1% |

Another interesting result is that one-year-old firms account for a disproportionate share of systematic churning, around 19% of it, while representing only 10% of the total number of surviving firms. This accounts for $-3\%$ of the within component, while their average growth rate accounts for 24% of the within component. This implies that overall one-year-old firms account for 21% of the within growth rate, affecting disproportionally both the average growth rate of productivity and churning.

# G Figures

# H Comparative Statics: Causes of Changing Business Dynamism

Here, I consider individually the effects of a decline in labor force growth and an increase in the fixed operating cost.

## H.1 Decline in Labor Force Growth

In this subsection, I discuss the effects of declining labor force growth. The decline in US labor force growth is well-known: labor force growth in the period 2000-2014 was 1 percentage point lower than in the period 1977 to 1999. The literature has already linked this slowdown to a reduction in entry (Hopenhayn et al., 2018; Karahan et al., 2019). This model can complement these

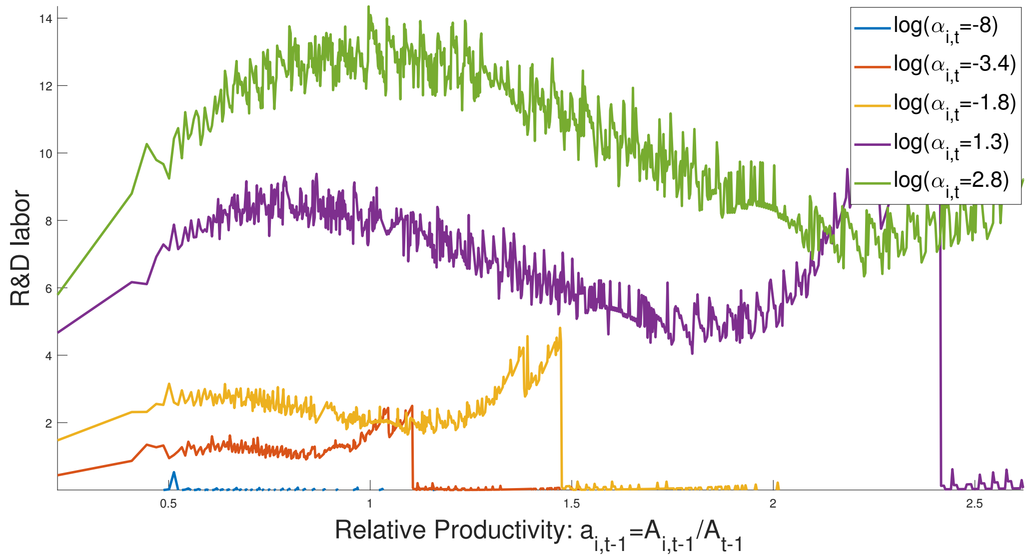


Figure 7: R&D labor as a function of relative productivity for selected levels of R&D productivity.

analyses by discussing the impact on the aggregate productivity growth rate.

Table 14 shows the impact of reducing labor force growth by 1 percentage point on selected moments. Unsurprisingly, this change explains the entire reduction in net entry. However, although the results are qualitatively in line with all changes in the dimensions of business dynamism, the quantitative effect is quite far.

The effect on aggregate productivity growth is negative for two reasons. First, lower labor force growth directly affects the semi-endogenous component of productivity growth: slower population growth implies slower steady state growth in the number of firms, thus slower growth coming from the expanding varieties channel. This is visible from proposition 1 and equation (31).

Second, a lower $\lambda$ reduces the aggregate productivity growth rate through lower R&D. This result is due to the reduction in the number of entrants, whose R&D productivity is on average higher than incumbents.

Overall, the reduction in labor force growth is responsible for reducing the entry rate by almost half of the data. Thus, despite not providing a comprehensive explanation for changing business dynamism, this channel is nevertheless relevant.

Table 14: Labor force growth declines by 1 percentage point

| Moment | Model | Data | |
|---|---|---|---|
| Δ Net entry rate | -1pp | -1pp | US Census |
| Δ Entry rate | -1.1pp | -2.4pp | US Census |
| Δ Reallocation rate | -0.5pp | -4pp | US Census |
| Δ Reallocation rate (cont.) | -0.1pp | -0.7pp | US Census |
| %Δ Lab. productivity growth | -4% | -53% (-1.5pp) | FRB SF ('94-'04 vs '05-'15 ) |
| | | -14% (-0.2pp) | FRB SF ('73-'93 vs '05-'15 ) |
| Δ R&D to GDP | -4% (-0.1pp) | +9%(+0.2pp) | NIPA ('94-'04 vs '05-'15 ) |
| | | +23% (+0.4pp) | NIPA ('77-'99 vs '00-'14) |
| %Δ Average plant size | +0.6% | +4.4% | US Census |

## H.2 Increasing Fixed Operating Cost

As Ghazi (2019) shows, the fixed operating cost has been rising over time, especially for large firms. Here I consider an increase of $\Phi$ by 10% to match the change in average establishment size. I summarize the effects in table 15.

The largest impact of this change is on the average establishment size, with a small effect on the increased entry rate. This change in entry results from an increase in the number of firms that find it unprofitable to stay in the market and decide to exit. Consequently, aggregate productivity growth increases because of a stronger cost-spreading effect that increases the incentive to innovate.

Table 15: Fixed operating cost $\Phi$ increases by 10% to match the change in average plant size.

| Moment | Model | Data | |
|---|---|---|---|
| Δ Net entry rate | 0pp | -1pp | US Census |
| Δ Entry rate | +0.2pp | -2.4pp | US Census |
| Δ Reallocation rate | +1.1pp | -4pp | US Census |
| Δ Reallocation rate (cont.) | +0.8pp | -0.7pp | US Census |
| %Δ Lab. productivity growth | +5% | -53% (-1.5pp) | FRB SF ('94-'04 vs '05-'15 ) |
| | | -14% (-0.2pp) | FRB SF ('73-'93 vs '05-'15 ) |
| Δ R&D to GDP | -5% (-0.1pp) | +9%(+0.2pp) | NIPA ('94-'04 vs '05-'15 ) |
| | | +23% (+0.4pp) | NIPA ('77-'99 vs '00-'14) |
| %Δ Average plant size (targeted) | +4.4% | +4.4% | US Census |