CFCM

CENTRE FOR FINANCE, CREDIT AND MACROECONOMICS

Working Paper 24/03

Hetereogeneous firms, growth and the long shadows of business cycles

- **Cristiana Bendetti-Fasil, Giammario Impullitti, Omar Licandro, Petr Sedláček and Adam Hal Spencer**

Produced By:

Centre for Finance, Credit and Macroeconomics
School of Economics
Sir Clive Granger Building
University of Nottingham
University Park
Nottingham
NG7 2RD

www.nottingham.ac.uk/cfcm

Hetereogeneous Firms, Growth and the Long Shadows of Business Cycles*

Cristiana Bendetti-Fasil Giammario Impullitti[†] Omar Licandro[‡]
Petr Sedláček[§] Adam Hal Spencer[¶]

August 5, 2024

Abstract

R&D is procyclical and a crucial driver of growth. Evidence indicates that innovation activity varies widely across firms. Is there heterogeneity in innovation cyclicality? Does innovation heterogeneity matter for business cycle propagation? We provide empirical evidence that more productive firms are less procyclical in innovation. We develop a model replicating this observation, with selection as the driver of heterogeneous innovation cyclicality. We then examine how heterogeneous innovation and growth influence business cycle propagation. Dynamics of firm entry and exit, coupled with heterogeneous cyclicality, significantly amplify TFP shock propagation. Business cycle fluctuations give substantial welfare losses, with firm heterogeneity contributing significantly.

Keywords: Growth, Business Cycles, Innovation, Heterogeneous Firms

*Our wonderful friend and colleague Cristiana Benedetti-Fasil is no longer with us. Finishing this project without her insights, energy and contagious enthusiasm has been hard and painful. A previous version of the paper was titled "Heterogeneous Firms, R&D Policies and the Long Shadows of Business Cycles".

[†]University of Nottingham, e-mail: giammario.impullitti@nottingham.ac.uk

[‡]University of Leicester & Barcelona Graduate School of Economics (GSE), e-mail: omar.licandro@leicester.ac.uk

[§]University of New South Wales, e-mail: p.sedlacek@unsw.edu.au

[¶]University of Nottingham, e-mail: adamhalspencer@gmail.com

1 Introduction

There is a long tradition in macroeconomics of studying endogenous growth and business cycles separately. However, events such as the financial crisis of 2009 and the recent pandemic have raised concerns that severe cyclical downturns may have persistent negative effects through their impact on long-run growth. Indeed, many countries have failed to revert back to their original growth paths following the Great Recession (Ball, 2014; Blanchard, 2015)[1]. Innovation and firm heterogeneity play an important role in shaping long-run growth patterns and the aggregate effects of macroeconomic shocks (Klette and Kortum, 2004; Acemoglu et al., 2018b; Sedláček and Sterk, 2017; Ates and Saffie, 2018). What is the role of innovation and heterogeneous firms for the transmission mechanisms and propagation of business cycle fluctuations?

To address this question, we develop a model of endogenous growth and business cycles featuring heterogeneous firms. Firms pay a sunk entry cost to draw a prototype efficiency from the support of the distribution of incumbent firms' productivity to produce a new variety of goods. To turn that prototype efficiency into effective productivity, produce and become incumbent firms they need to successfully innovate. Once a firm becomes an incumbent it keeps innovating to improve its productivity and survive on the market. A stochastic fixed operating cost introduces a selection mechanism, causing firms to exit the market if the cost becomes prohibitively high, relative to their productivity and hence profitability. Knowledge spillovers sustain long-run growth driven by innovation from both incumbents and entrants, while short-run fluctuations are generated by aggregate, mean-reverting productivity shocks.

Calibrating the model to match salient firm-level and aggregate statistics of the US economy, we analyze the role of innovation and firm heterogeneity in shaping the propagation of business cycle shocks. Besides the well-known, and empirically grounded, prediction that innovation is procyclical[2], the model offers a novel prediction that innovation is more procyclical for smaller, less productive, less profitable, firms. This is a key feature of our theory, which highlights the importance of introducing firm heterogeneity in models of business cycles and growth. To validate this prediction we empirically test it using Compustat data, which indeed reveal that R&D spending is more procyclical for smaller, less profitable firms.

[1] The observation that recoveries may not always be robust enough to return GDP to its pre-crisis trend extends beyond this particular crisis. Empirical evidence abundantly shows that GDP fluctuations are persistent, with their effects lingering for years after the initial shock. See Cerra et al. (2023) for a recent survey.

[2] Empirical evidence on the procyclicality of aggregate R&D can be found in Griliches (1990); Comin and Gertler (2006); Barlevy (2007); Ouyang (2011); Aghion et al. (2012) among others.

The sources of aggregate R&D cyclicality and its heterogeneity across firms in our economy can be traced to several key factors. Our model, like most endogenous growth models, features intertemporal knowledge spillovers: when a firm innovates, it produces new ideas that future firms can build upon, creating a positive externality. This leads firms to underinvest in innovation from a societal perspective and introduces a present-bias in innovation. An innovating firm, aware that due to knowledge spillovers, subsequent innovations will slowly reduce its market shares to replace it, prioritizes short-term profitability over long-term benefits. As a result, entrepreneurs become shortsighted and more sensitive to short-term profit fluctuations. Selection mechanisms drive the relatively higher innovation cyclicality of less productive firms. The risk of exit is greater for these firms, prompting them to respond more aggressively to profit fluctuations caused by cyclical shocks. This dynamic explains why less productive firms exhibit higher innovation cyclicality in the face of economic fluctuations.

Counterfactual analysis reveals that heterogeneity in innovation among both incumbents and entrants substantially contributes to the transmission of TFP shocks. We perform two counterfactuals that sequentially shut down the key dimensions related to firm heterogeneity. First, we impose common innovation levels and cyclicality across firms, effectively eliminating selection. While entry and exit still occur, the innovation success probability no longer depends on firms' productivity. The second counterfactual shuts down endogenous entry and exit, ensuring that the mass of firms and the survival probability do not respond to shocks. Our findings indicate that both layers of heterogeneity lead to greater volatility and amplification of shocks. A 10% drop in GDP, roughly corresponding to the one produced by the COVID shock, generates a persistent recession. Importantly, even twenty years after the shock hits, GDP in our baseline economy remains 4% below what it would have been under trend (steady-state) growth. Firm heterogeneity accounts for approximately 1/4 of this difference. In our framework, procyclical entry and countercyclical exit of less productive or younger firms can potentially mitigate the impact of technology shocks, akin to the "cleansing effects" of recessions described by Caballero and Hammour (1994). However, our novel insight is that these firms also exhibit the most procyclical innovation behavior, which amplifies rather than dampens the effects of business cycles.

Intuitively, in the first counterfactual, firm entry and exit respond to the shock, affecting the mass of firms and reallocating market shares among them. This reallocation occurs across firms with identical innovation levels and cyclicality, unlike in the baseline model where these are heterogeneous. In our baseline economy, a boom increases the mass of less productive and younger firms, which are more procyclical. This reallocation amplifies the size and persistence of business cycle fluctuations. When entry and exit do not respond to the shock, as in the second counterfactual, its propagation is further weakened, as there is no reallocation at all.

In his seminal work, Lucas (1987) demonstrated that the cost of economic fluctuations is negligible in most standard business cycle models, amounting to less than 0.1% of annual consumption. Endogenous productivity growth typically increases these costs. What is the role of firm heterogeneity in shaping the cost of economic fluctuations? As in Barlevy (2004), we find that endogenous growth substantially raises the costs of business cycle fluctuations. Importantly, the presence of firm-level heterogeneity in our endogenous growth model exacerbates such business cycle costs by an additional 15%.

Literature review. The paper is related to several lines of research in the literature. First, our work is related to the recent literature that bridges endogenous growth and the business cycle approaches to macroeconomic analysis.[3] A key challenge that all models linking business cycle and endogenous growth have to face is to reproduce the pro-cyclicality of innovation observed in the data. The theoretical underpinning of pro-cyclical innovation is also fundamental in shaping the role of growth and in particular of technical change for the propagation and persistence of short-run fluctuations. Aghion et al. (2010) introduce credit constraints to offset the opportunity cost channel linking innovation and business cycle. A negative productivity shock reduces current cash flow, which in turn reduces firm capacity to borrow to finance investment improving productivity in the future.[4]

Another explanation for the pro-cyclicality of R&D is endogenous labor supply (Fatas, 2000). If labor supply is pro-cyclical, R&D will be pro-cyclical as well, since firms will be reluctant to divert already scarce resources away from production. Pro-cyclicality of R&D can also occur if innovation is produced with goods, not labor. During booms, when more goods are produced, there are more resources available for R&D. Aghion and Saint-Paul (1998) and Comin and Gertler (2006) follow this route.

As in Barlevy (2007), our model posits that the pro-cyclicality of R&D stems from the heavy dependence of R&D incentives on the short-run profits of innovating firms. This dependence dominates the "opportunity cost" channel, which suggests that firms should innovate more during downturns because the foregone output and sales are lower (Aghion and Saint-Paul, 1998; Aghion et al., 2010). Barlevy (2007) shows that if production fixed costs are large, profits are strongly pro-cyclical, and so is R&D. Specifically, free entry

[3] See Fatas (2000), Barlevy (2004, 2007), Comin and Gertler (2006), Aghion et al. (2009, 2010, 2014), Nuno (2011), Anzoategui et al. (forthcoming), Bianchi et al. (2019), Benigno and Fornaro (2018), Vinci and Licandro (2020),Cozzi et al. (2021) and Fornaro and Wolf (2023), among others.

[4] A similar argument is presented in Stiglitz (1994), where imperfections in financial markets serve as the crucial link between business cycle fluctuations, innovation and growth.

requires that for innovation incentives to be larger during booms, profits must be more pro-cyclical than the cost of R&D.

Firm heterogeneity amplifies the pro-cyclical forces in our model, as fixed costs imply that low-productivity firms are more sensitive to profit fluctuations due to their higher likelihood of exit. Moreover, if marginal firms engage in some R&D, their exit reduces the aggregate innovation effort. Thus, firm heterogeneity adds new channels that strengthen the pro-cyclicality of innovation, thereby enhancing the propagation and persistence of short-run fluctuations.

Our paper is also related to the strand of literature which investigates the role of heterogeneous firms in aggregate fluctuations (see e.g. Caballero and Hammour, 1994; Ghironi and Melitz, 2005; Ottaviano, 2012; Moscarini and Postel-Vinay, 2012; Lee and Mukoyama, 2015; Boucekkine et al., 1997, 2005). While existing studies primarily focus on differences across firms in terms of job creation, exit and entry of heterogeonous producers we pay particular attention to heterogeneity in their innovative efforts. Ottaviano (2012) introduces heterogeneity into a standard business cycle model with entry and exit to analyze the effect of technology shocks. The model generates a pattern of procyclical entry and countercyclical exit that dampens the overall effect of technology shocks on aggregate output and welfare, as new entrants and exiting firms are generally less efficient than established firms. Bilbiie et al. (2012) perform a similar exercise but abstracting from firm heterogeneity and find that entry and exit can substantially enhance the propagation of business cycles. Ottaviano (2012) concludes that using representative firms models may lead to overstate the role of procyclical entry and exit as a propagation mechanism of technology shocks.

There is neither innovation nor long-run growth in the economies analyzed in the studies discussed above. We depart from them by introducing innovation-driven growth and, more importantly, by assessing the importance of heterogeneity in the innovation response to business cycle fluctuations of both entrants and incumbent firms.[5] Differently from Ottaviano (2012) we find that entry and exit of heterogeneous innovators magnifies the effects of technology shocks on output and welfare. Thus this suggests that the results in Bilbiie et al. (2012) can still hold in economies with heterogeneous firms if we introduce within-firm, innovation-driven productivity dynamics. Intuitively, while procyclical entry and countercyclical exit can potentially dampen the propagation of fluctuations as they involve less productive (and younger) firms, these firms also feature the most procyclical innovation

[5] As such, our paper is related to Ates and Saffie (2018) and Schmitz (2021) who study the heterogeneous response of firms' innovation to financial crises and Sedláček (2020) who study the impact of faster growth on uncertainty. In contrast to these papers, we highlight the differences in cyclical behavior of firms' innovation and its role for overall aggregate fluctuations and growth.

which has opposite effect.

Finally, we make contact with the literature on the welfare cost of business cycle. Lucas (1987) sets up the framework through which to evaluate the welfare costs of business cycle fluctuations and concluded that – relative to the benefits of even small changes in an economy's growth rate – they are very small. Following Lucas' analysis, there have been numerous attempts at developing models in which business cycles were costly (see Lucas, 2003, for a summary). One line of work has considered preferences which make fluctuations more painful (e.g. Tallarini, 2000). Other papers have focused on individual heterogeneity in the exposure to business cycles (e.g. Krusell et al., 2009; Krebs, 2007). Another strand of literature has put forward the notion that business cycles may in fact change aggregate growth, thereby having larger impact on welfare (e.g. Krebs, 2003; Barlevy, 2004). Our paper is closely related to this approach as it links business cycle fluctuations to aggregate growth. We contribute by showing that firm heterogeneity plays a crucial role in the quantitative impact of business cycles on growth and welfare.

The rest of the paper is structured as follows. The next section introduces the baseline model. Section 3 describes the parametrization. Section 4 provides the main quantitative analysis of the baseline and analyze the implication of heterogeneous innovation for the cost of business cycle. Section 5 concludes.

2 Model

In this section, we describe our theoretical model. To ease the exposition, we use upper-case letters for aggregates and lower-case letters for firm-level variables.

2.1 Environment

The model is in discrete time $t \in \{0, 1, 2, ...\}$ and has three types of agents — intermediate goods firms, a representative household and final goods producer.

Household. The representative household inelastically supplies labor, N, to intermediate goods producers and chooses consumption, C, and asset holdings, A, to maximize life-time utility

$$\max_{\{C_t, A_{t+1}\}} \quad \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma}}{1-\sigma} \tag{1}$$

where $\mathbb{E}$ is an expectations operator, $\beta \in [0,1]$ is the discount factor and σ is the coefficient of relative risk aversion. The household makes its choices subject to the following budget constraint constraint

$$P_t C_t + A_{t+1} = W_t N + (1 + R_t) A_t \tag{2}$$

where P_t is the aggregate price index, W_t is the wage rate and R_t is the return on assets.

Final Goods Firms. We assume that the final goods firm combines intermediate goods according to the following production function

$$Y_t = \left(\int_{j \in \Omega_t} y_t(j)^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \tag{3}$$

where $y_t(j)$ denotes the quantity of intermediate good j at time t, Ω_t is both the measure and set of incumbent intermediate firms and $\eta > 1$ is the elasticity of substitution between varieties. These final goods producers are assumed perfectly competitive, taking the CPI P_t as given.

Incumbent Intermediate Goods Firms. Each variety is produced with the objective of maximizing the expected present value of dividends to the firm's equityholders using production function

$$y_t(j) = Z_t q_t(j) n_t(j), \tag{4}$$

where $n_t(j)$ denotes labor resources utilized, $q_t(j)$ is *endogenous* firm-level productivity and Z_t is exogenous aggregate total factor productivity. The latter is assumed to evolve according to

$$\log(Z_t) = \rho_Z \log(Z_{t-1}) + \epsilon_{Z,t} \tag{5}$$

where $\epsilon_{Z,t} \sim N(0, \sigma^2_{Z,t})$ is a stochastic shock. In contrast, the evolution of firm-level productivity $q_t(j)$ is endogenous, driven by firms' investment into research and development (R&D). In particular, we assume that using $s_t(j)$ units of the final good delivers an $x_t(j) \in [0,1]$ probability of innovation according to

$$s_t(j) = \gamma \widehat{q}_t(j)^{\eta-1} x_t(j)^{\psi} \tag{6}$$

where $\gamma > 0$ is the inverse of the efficiency of R&D, $\psi > 1$ is a curvature parameter making the R&D cost function convex, $\widehat{q}_t(j) = q_t(j)/Q_t$ is normalized firm-level productivity with Q_t being the average productivity index defined as

$$Q_t = \left(\frac{1}{\Omega_t} \int_{j \in \Omega_t} q_t(j)^{\eta-1} dj \right)^{\frac{1}{\eta-1}}. \tag{7}$$

Notice that $\eta - 1 > 0$ measures two things at the same time. Firstly, the negative effect of firm-specific productivity on firm-specific R&D efficiency. Secondly, the the positive knowledge spillovers from average productivity. Firms with a firm-specific productivity below average are more efficient in R&D than the average firm. Notice also that parameter η governs the degree of R&D heterogeneity between incumbent firms.

Using the definition of (7), we define the endogenous growth rate of the economy as $G_t = Q_t/Q_{t-1} - 1$ (to elaborated on later). In the event of a successful innovation, the firm's idiosyncratic productivity improves by a factor of $\lambda > 0$ in the next period. If innovation efforts fail, firms' productivity levels remain unchanged. Therefore, idiosyncratic firm productivity evolves according to

$$q_{t+1}(j) = \begin{cases} q_t(j)(1+\lambda) & \text{with probability } x_t(j) \\ q_t(j) & \text{with probability } 1 - x_t(j). \end{cases} \tag{8}$$

Incumbent firms need to pay fixed costs, ϕ in order to stay in operation. We assume that fixed costs are stochastic, distributed independently and identically over time and across firms according to a distribution $F(\phi)$. Finally, to obtain a realistic firm-size distribution, similarly to Sedláček (2020), we assume that firms realize deterministic efficiency gains, through learning-by-doing as they age. Specifically, we assume firms of age group $a_t \in \{1, 2, .., n_A\}$, experience exogenous efficiency gains proportional to their own employment level $n_t(j)$. Efficiency gains reduce the labor cost $W_t n_t(j)$ by $Q_t \psi_a n_t(j)$.[6] A firm of age group $a_t < n_A$ moves to the next age group in the following period $a_{t+1} = a_t + 1$ with

[6] As shown below, this assumption ensures that the firm's wage bill will for the period be $(W_t - Q_t \psi_a) n_t(j)$.

probability ζ_a. Firms in age group n_A remain there unless exiting the industry. We also allow for an exogenous probability of firm exit, which varies with age group, denoted by δ_a.

Entrant Intermediate Goods Firms. Entrants incur sunk cost $c_e > 0$ at establishment in t. Subsequently, they develop a prototype variety, whose idiosyncratic productivity $q_{e,t}$ is drawn from a normal distribution truncated in the support of the incumbents' productivity with parameters μ_e for position and σ_e for dispersion. We denote this distribution as $\omega_{e,t}(q_{e,t})$, where index t indicates that this support is changing over time. After drawing their prototype productivity, entrants make an R&D investment choice, using similar technology to incumbents

$$s_{e,t}(j) = \gamma_e \widehat{q}_{e,t}(j)^{\eta-1} x_{e,t}(j)^{\psi} \tag{9}$$

where $s_{e,t}$ denotes their final goods input at time t, γ_e is productivity of their innovation technology, $\widehat{q}_{e,t} = q_{e,t}/Q_t$ is their prototype relative productivity and $x_{e,t} \in [0,1]$ is their probability of successful innovation. We assume that these potential entrants must successfully innovate upon their prototype in order to become incumbents at time $t+1$. In such event, they commence operations at $t+1$ with productivity of $(1+\lambda)q_{e,t}$, in the first age group. Should they fail to successfully innovate (with probability $1 - x_{e,t}$), they exit the industry and their prototypes become obsolete. We denote the mass of potential entrants (prior to the outcome of their innovation being realized) by M_t.

Timing. We summarize the period t timing for incumbents and entrants. The period t timing for incumbent intermediate firms is

1. Draw *iid* fixed cost shock ϕ_t and make endogenous exit choice.
2. Conditional on remaining, make optimal static choice of n_t and produce with technology (4).
3. Make optimal inter-temporal R&D choice s_t, x_t and distribute dividends to households.
4. Draw age group and death shocks.
5. Conditional on survival, draw innovation success shock from distribution Bernoulli(x_t) and move to time $t+1$.

Then for new entrants at time t, the timing is

i. Pay sunk cost c_E and draw prototype productivity level.

ii. Make optimal choice of R&D investment $s_{e,t}$ and $x_{e,t}$.

iii. Draw innovation success shock from Bernoulli$(x_{e,t})$.

iv. If successful in innovating, at time $t+1$, proceed to Step 1 of the incumbent timing above. Otherwise, exit.

2.2 Equilibrium

Here we detail the optimal choices of agents, given the setup in Section 2.1. We take the final good to be the numéraire, meaning that $P_t = 1\,\forall t$.

Household. The household's optimal savings decision results in the Euler equation

$$1 = \beta \mathbb{E}_t \left[(1 + R_{t+1}) \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \right], \tag{10}$$

from which we define the stochastic discount factor (SDF) $\Lambda_t = \beta \left(C_{t+1}/C_t \right)^{-\sigma}$ to use in the intermediate firms' problem.

Final Goods Firm. The final goods firm chooses its purchase of each variety to maximize its profits, subject to aggregation technology (3), giving variety-level demand

$$y_t(j) = p_j(t)^{-\eta} Y_t, \tag{11}$$

as well as the aggregate price index aggregate price index

$$P_t = \left(\int_{j \in \Omega_t} p_t(j)^{1-\eta} dj \right)^{\frac{1}{1-\eta}}. \tag{12}$$

Incumbent Intermediate Goods Firms. The firms' state vector is comprised of their idiosyncratic relative productivity and age, as well as aggregate TFP. From now onwards, we replace variety-level notation with a firm's state; also dropping time subscripts and adopting the convention that variables with $'$ superscripts are for time $t+1$. The firms' state is thus represented as $(\widehat{q}, a, Z)$. The firms' recursive formulation at the beginning of the period, conditional on remaining in operation, is given as

$$v(\widehat{q},a,Z)=\max_{\{p,n,x\}}\ \pi(\widehat{q},a,Z)+\mathbb{E}\Lambda'\left[x\widetilde{v}\left(\frac{\widehat{q}(1+\lambda)}{1+G},a',Z'\right)+(1-x)\widetilde{v}\left(\frac{\widehat{q}}{1+G},a',Z'\right)\right] \tag{13}$$

subject to

$$\pi(\widehat{q},a,Z)=p\widehat{y}-(\widehat{W}-\psi_a)n-s \tag{14}$$

$$\widetilde{v}(\widehat{q}',a',Z')=\int_\phi \max\left[v(\widehat{q}',a',Z')-\phi,0\right]dF(\phi) \tag{15}$$

and the form of the SDF in (10), the demand curve (11), innovation production function (6), production function (4) and TFP process (5). Equation (13) states that firms choose their price, employment and R&D investment to maximize the sum of period profits and expected continuation value.[7] See that the expectation operator takes account of the aggregate TFP process, as well as the idiosyncratic ageing shock for next period's a' and the exogenous death shock. Equation (14) gives profits in their de-trended form, where $\widehat{y}$ and $\widehat{W}$ are firm-level output and wages, relative to aggregate Q. Equation (15) accounts for firms' endogenous exit choices, where they choose to receive their continuation value less than the burden of fixed costs. The resulting optimal decisions are then given by

$$p(\widehat{q},a,Z)=\frac{\eta}{\eta-1}\frac{\widehat{W}-\psi_a}{Z\widehat{q}}, \tag{16}$$

$$n(\widehat{q},a,Z)=\frac{\widehat{y}}{\widehat{q}}, \tag{17}$$

$$\widetilde{\phi}(\widehat{q},a,Z)=\widetilde{v}(\widehat{q},a,Z), \tag{18}$$

$$x(\widehat{q},a,Z)=\left[\mathbb{E}\frac{\Lambda'\left\{\widetilde{v}\left(\frac{\widehat{q}(1+\lambda)}{1+G},a',Z'\right)-\widetilde{v}\left(\frac{\widehat{q}}{1+G},a',Z'\right)\right\}}{\gamma\psi\widehat{q}^{\eta-1}}\right]^{\frac{1}{\psi-1}}. \tag{19}$$

The above optimality conditions show that firms choose prices as a constant markup over marginal costs (16), which relies on age-dependent efficiency. They employ workers in order to satisfy the demand constraint (17). Firms shut-down when their expected value falls below a cutoff operational cost, which depends on their state $\widetilde{\phi}(\widehat{q},a,Z)$, as given by (18).[8]

[7] Notice that we express everything as a function of the firm's detrended productivity. As such the future period's productivity argument declines at the rate of productivity growth in the case of innovative failure, giving $\widehat{q}/(1+G)$, while jumping at the step size to give $\widehat{q}(1+\lambda)/(1+G)$ in the case of success.

[8] Using this definition, we can express the value function (15) as $\widetilde{v}(\widehat{q}',a',Z') = F(\widetilde{\phi}(\widehat{q}',a',Z'))\left[v(\widehat{q}',a',Z')-\mathbb{E}[\phi|\phi<\widetilde{\phi}(\widehat{q}',a',Z')]\right]$.

They lastly choose R&D investment to balance the marginal cost and benefits of innovating (19). The latter is given precisely by the difference in firm value conditional on innovating and not, as in the numerator of (19).

Entrant Intermediate Goods Firms. Entrants make a choice of R&D investment, conditional on their prototype draw, to maximize conditional value

$$v_e(\widehat{q}_e, Z) = \max_{\{x_e\}} -s_e + \mathbb{E}\Lambda' x_e \widetilde{v}\left(\frac{\widehat{q}_e(1+\lambda)}{1+G}, 1, Z'\right) \tag{20}$$

where the expression for the continuation value comes from that of an incumbent in the first age group in (15). Entrants maximize (20) subject to innovation production function (9), yielding FOC

$$x_e = \left(\mathbb{E}\frac{\Lambda' \widetilde{v}\left(\frac{\widehat{q}_e(1+\lambda)}{1+G}, 1, Z'\right)}{\widehat{q}_e^{\eta-1}\psi\gamma_e}\right)^{\frac{1}{\psi-1}}. \tag{21}$$

Comparing the right-hand side of (21) with that of the incumbents (19) highlights the assumption regarding the survival of entrants with respect to innovation. Entrants must innovate to survive, hence their marginal benefit of investment is the absolute firm value of successful innovation $\widetilde{v}\,(\widehat{q}_e(1+\lambda)/1+G, 1, Z')$, rather than the increment in the case of incumbents $\widetilde{v}\,(\widehat{q}(1+\lambda)/1+G, a', Z') - \widetilde{v}\,(\widehat{q}/1+G, a', Z')$. We assume that a free entry condition holds in steady state, specifically $\Delta = 0$ where $\Delta = \int_{\widehat{q}_e} v_e(\widehat{q}_e, 1)d\omega_{t,e}(\widehat{q}_e) - Wc_E$, which determines the deterministic steady state entry mass $\bar{M}$ and Wc_E is spending on the sunk entry cost in terms of labor. However, along the transition after a TFP shock, we instead allow a for a more general equation to determine the entry mass

$$M = \bar{M}\exp(\chi\Delta) \tag{22}$$

where $\chi \geq 0$ is a parameter governing the elasticity of entry with respect to firm value. This parameter allows us to control the strength of the entry response to an aggregate TFP shock. Note that when $\chi = 0$, entry remains fixed at its steady state value along the transition; instead when $\chi \to \infty$, free entry holds at every instant.

Equilibrium Definition. Here we give an abridged version of the equilibrium definition. More details are deferred to Appendix A.2. A stochastic stationary equilibrium is such that

1. All agents are optimizing (households, final goods firms, intermediate goods firms).

2. All markets are clearing. For final goods

$$C + S + \Phi - \Psi = Y \tag{23}$$

where S is R&D expenditure by both incumbents and potential entrants, Φ represents aggregate fixed operating costs, Ψ are aggregate age-related efficiency gains and Y is as in (3). Labor market clearing is given by

$$E + L = N \tag{24}$$

where $E = Mc_E$ represents aggregate entry costs, L is aggregate variable labor demand and N is the fixed endowment of labor from the household.

3. There exists a cross-sectional measure of firms over states, which evolves endogenously. Denote this measure of firms over relative productivity $\widehat{q}$, over age groups a, with aggregate state Z by $\omega(\widehat{q}, a, Z)$. Given that new entrants start in the first age group, we need to consider it separately from those age groups that follow. The cross-section's law of motion for $a > 1$ is given by

$$\begin{aligned}
\omega'(\widehat{q}, a, Z') = {} & (1-\zeta_a)(1-\delta_a) \int^{\widetilde{\phi}(\widehat{q},a,Z')} x\left(\frac{1+G}{1+\lambda}\widehat{q}, a, Z\right) \omega\left(\frac{1+G}{1+\lambda}\widehat{q}, a, Z\right) \mathrm{d}F(\phi) \\
& + (1-\zeta_a)(1-\delta_a) \int^{\widetilde{\phi}(q,a,Z')} \Big(1 - x\big((1+G)\widehat{q}, a, Z\big)\Big) \omega\big((1+G)\widehat{q}, a, Z\big) \mathrm{d}F(\phi) \\
& + \zeta_{a-1}(1-\delta_{a-1}) \int^{\widetilde{\phi}(\widehat{q},a,Z')} x\left(\frac{1+G}{1+\lambda}\widehat{q}, a-1, Z\right) \omega\left(\frac{1+G}{1+\lambda}\widehat{q}, a-1, Z\right) \mathrm{d}F(\phi) \\
& + \zeta_{a-1}(1-\delta_{a-1}) \int^{\widetilde{\phi}(q,a,Z')} \Big(1 - x\big((1+G)\widehat{q}, a-1, Z\big)\Big) \omega\big((1+G)\widehat{q}, a-1, Z\big) \mathrm{d}F(\phi)
\end{aligned} \tag{25}$$

where note that $\zeta_a = 0$ for the highest age group $a = n_A$. The first term on the right-side of (25) captures incumbent firms who are successful in innovating and who remain in age grouping a. Such firms move upwards in their relative productivity at net rate of λ, while also drifting downwards at the growth rate of $Q(t)$. The second term captures incumbents who are unsuccessful in innovating, who again remain in age group a. The third and fourth terms pertain to incumbents who age, moving up to group a from $a - 1$, who are successful and unsuccessful in innovating, respectively.

The first age group, $a = 1$, also accounts for new entrants, as specified below

$$\omega'(\widehat{q}, a, Z') = (1-\zeta_a)(1-\delta_a) \int^{\widetilde{\phi}(\widehat{q},a,Z')} x\left(\frac{1+G}{1+\lambda}\widehat{q}, a, Z\right) \omega\left(\frac{1+G}{1+\lambda}\widehat{q}, a, Z\right) \mathrm{d}F(\phi)$$
$$+ (1-\zeta_a)(1-\delta_a) \int^{\widetilde{\phi}(q,a,Z')} \Big(1 - x\big((1+G)\widehat{q}, a, Z\big)\Big) \omega\big((1+G)\widehat{q}, a, Z\big) \mathrm{d}F(\phi)$$
$$+ M \int^{\widetilde{\phi}(\widehat{q},1,Z')} x_e\left(\frac{1+G}{1+\lambda}\widehat{q}, Z\right) \omega_e\left(\frac{1+G}{1+\lambda}\widehat{q}\right) \mathrm{d}F(\phi). \quad (26)$$

The first and second terms in (26) are for incumbents who remain in the first age grouping, who were successful and unsuccessful in innovating, respectively. The final term represents new entrants — recall such firms only become active incumbents if they successfully innovate over their prototype draw.

4. The mass of entrants is pinned down by (22).

5. The equilibrium growth rate is determined by $G' = Q'/Q - 1$ where Q is given by (7). See Appendix A.1 for the full expression of this object. We present the definition of the growth rate below for a special case of the model with one single age grouping ($n_A = 1$), no exogenous exit ($\delta_1 = 0$) and in the deterministic steady state ($Z = Z' = 1$).[9]

$$\Omega'(1+G')^{\eta-1} = \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}(1+\lambda)/(1+G'))} [\widehat{q}(1+\lambda)]^{\eta-1} x(\widehat{q}) \omega(\widehat{q}) \mathrm{d}F(\phi) d\widehat{q}$$
$$+ \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}/(1+G'))} \widehat{q}^{\eta-1} [1 - x(\widehat{q})] \omega(\widehat{q}) \mathrm{d}F(\phi) d\widehat{q}$$
$$+ M \int_{\widehat{q}_e} \int^{\widetilde{\phi}(\widehat{q}_e(1+\lambda)/(1+G'))} \widehat{q}_e(1+\lambda) x_e(\widehat{q}_e) \omega_e(\widehat{q}_e) \mathrm{d}F(\phi) d\widehat{q}_e. \quad (27)$$

The first two terms of (27) pertain to incumbent firms, while the last relates to entrants. An incumbent firm can successfully innovate (first set of integrals); its relative productivity will rise to $\widehat{q}(1+\lambda)/(1+G')$, which forms the basis of its exit choice $\widetilde{\phi}(\widehat{q}(1+\lambda)/(1+G'))$, given the model timing. These firms innovate successfully with probability $x(\widehat{q})$ and their measure is given by $\omega(\widehat{q})$. The second term relates to incumbent firms that are unsuccessful in innovating. This happens with probability $[1-x(\widehat{q})]$, while their exit cutoff is based-on their relative productivity, which drifts downwards relative to the average, $\widetilde{\phi}(\widehat{q}/(1+G'))$. New entrants only contribute to growth when

[9] Note that, in this special case, we simplify the notation by dropping age and TFP indexing from the state space.

Table 1: Parameter values

Parameter		Value	Source/Target
Discount factor	β	0.96	Annual interest rate 4%
Coefficient of relative risk aversion	σ	1.00	Logarithmic preferences
TFP persistence	ρ_Z	0.78	Sedláček and Sterk (2017)
Innovation function curvature	ψ	2.00	Hall et al. (2001); Bloom et al. (2002)
TFP volatility	σ_Z	0.02	Real GDP volatility
Fixed cost mean	μ_ϕ	-2.62	Overall exit rate
Fixed cost variance	σ_ϕ	1.69	Entrant exit rate
Elasticity of substitution	η	2.05	Profits/GDP ratio
Innovation step size	λ	0.04	Growth rate
Incumbent innovation productivity	γ	0.95	Aggregate R&D to GDP
Entrant prototype mean	μ_e	0.00	Normalisation
Entrant prototype variance	σ_e	0.08	Mean prod. entrants/incumbents
Entrant innovation productivity	γ_e	1.05	Entrant contribution to growth
Elasticity of entry	χ	200.0	All positive R&D response
Exogenous death rate $a = 1$	δ_1	0.07	Exit rate firms aged 1–5
Exogenous death rate $a = 2$	δ_2	0.01	Exit rate firms aged 6–10
Exogenous death rate $a = 3$	δ_3	0.00	Normalisation
Efficiency gains $a = 1$	ψ_1	-0.46	Normalisation
Efficiency gains $a = 2$	ψ_2	-0.21	Size premium age 6–10 versus 1–5
Efficiency gains $a = 3$	ψ_3	0.00	Size premium age 11+ versus 1–5

they are successful in innovating, where their double integral takes similar form to successful incumbents; their contribution is weighted by their measure M.

3 Calibration

In this section, we detail the parameter values used in our quantitative analysis. We calibrate the deterministic steady state of the model to reproduce data facts for the U.S. economy. We defer the description of the solution methods to Appendix B. All parameter values are listed in Table 1.

Standard Choices. The following parameters are set to “standard” values in the literature. In particular, we set the discount factor, β, to 0.96 reflecting a roughly 4 percent interest rate at the annual frequency of our model. We assume logarithmic preferences with $\sigma = 1$. We take the persistence of the TFP process from Sedláček and Sterk (2017). Lastly, we assume that R&D costs are quadratic in the innovation rate, conforming with estimates

Table 2: Moments

Moment	Data	Model	Source
Volatility of real GDP	0.02	0.02	BEA
Volatility of consumption*	0.02	0.02	BEA
Volatility of R&D*	0.03	0.04	BEA
Overall Exit rate	0.11	0.11	BDS
Entrant exit rate	0.24	0.11	BDS
Profits/GDP ratio	0.12	0.14	BEA
Productivity growth	0.82	0.62	BLS
Aggregate R&D to GDP	0.04	0.06	BEA
Mean prod. entrants/incumbents	1.02	1.04	Foster et al. (2006)
Entrant contribution to growth	0.39	0.32	Pancost and Yeh (2022)
Exit rate firms age 1–5	0.16	0.17	BDS
Exit rate firms age 6–10	0.08	0.09	BDS
Size premium firms age 6–10	1.51	1.45	BDS
Size premium fimrs age 11+	2.06	2.02	BDS

Note: moments targeted, unless denoted with *. All numbers are prior to multiplication by 100, except for the average productivity growth: this is presented as a percentage after x100.

in the data (see e.g. Hall et al., 2001; Bloom et al., 2002).

Indirect Inference. The remaining parameters are set using an indirect inference approach along the lines of Acemoglu et al. (2018a). In particular, we compute selected model-generated moments and compare them to their respective counterparts in the data by minimizing

$$\min \sum_u \left(\frac{\text{model(u)} - \text{data(u)}}{\text{data(u)}} \right)^2,$$

where u indicates a given moment. While in general, all parameters affect the model's performance, in what follows we discuss them in relation to moments in the data to which they are most closely linked. In computing our empirical moments, we rely on data spanning 1987–2019. The selected moments and their empirical and model-implied values are presented in Table 2.

Three parameters are closely related to firm-level, and therefore aggregate, growth: innovation step size, λ and R&D efficiencies, γ for incumbents and γ_e for entrants. To parameterize these, we match the observed 0.82% average growth rate of TFP taken from the Bureau of Labor Statistics (BLS), 4% total R&D expenditures to GDP from the Bureau of Economic Analysis (BEA) and the overall contribution of entrants to growth of around 40%, taken from Pancost and Yeh (2022).[10]

The distribution of operational costs governs firms' exit patterns. Towards this end, we assume that the distribution is log-normal with mean and variance given by μ_ϕ and σ_ϕ^2, respectively. We set these parameters such that our model matches the observed 10.9% average firm exit rate from the Business Dynamics Statistics of the Census Bureau and the 23.7% exit rate for new entrants.

The volatility of aggregate productivity shocks is set so that our model matches the standard deviation of real GDP at business cycle frequencies in the data. Entrants' productivity is assumed to be drawn from a log-normal distribution with mean normalized to 0 and standard deviation, σ_e, targets the average relative productivity of entrants to incumbents of 1.02 reported in Foster, Haltiwanger and Krizan (2006). We target the overall profit to GDP ratio to pin-down the elasticity of substitution across varieties η, given that this parameter maps into the markup of firms set in the model. Profits in our model include R&D expenditures, so we combine the BEA pre-tax profit share of income with the R&D to GDP ratio to obtain a data target of 12.8% in our period of analysis. We set the elasticity of entry iteratively along the transition, to ensure that all incumbent firms' R&D responds positively on impact to a business cycle shock.[11]

We parameterize the firm age-size distribution such that there are three age groupings. Firms that are young (age 1–5 in the data), medium (age 6–10 in the data) and old (age 11+ in the data). To match the transitions across groupings young to medium and medium to old, we set the exogenous ageing probabilities, ζ_1 and ζ_2, to be 1/5, meaning firms take 5 years to age on average, (via a geometric distribution). We adjust the exogenous death probabilities across age groups to match their corresponding exit rates in the data. Given old firms have the lowest exit rate in the data, we normalize their exogenous death rate to zero. We target the relative sizes, in terms of employees, of medium and old firms to young. Here we normalize the exogenous efficiency gains to zero for old firms, giving negative values of the exogenous gains for young and medium-aged firms as in Sedláček (2020).

Table 2 shows that the model provides a good match of the empirical targets related to growth and firm heterogeneity. It also matches quite well the volatility of aggregate variables in relation to business cycles. We specifically look at that of real GDP, consumption and aggregate R&D spending. In generating these moments in the data, we take the natural logarithm of the corresponding time series and then de-trend using the HP filter with a smoothing

[10] Our model analogue of this contribution involves computing a counterfactual growth rate, where the entrants that become incumbents do so without experiencing a step increment over their prototype draw.

[11] This parameter translates the R&D impact responses of incumbents vertically. For more details, see Appendix E.

parameter of 100. The model performs well in matching all three volatility statistics, despite those for consumption and R&D being untargeted in the calibration exercise.

4 Quantitative analysis

In this section we present our main results. We begin by documenting properties in relation to firm heterogeneity — for both the steady state and along the transition. We then move on to its growth implications.

4.1 Firm Heterogeneity

We begin describing the extent of firm heterogeneity in our model. We do so by first concentrating on the stationary steady state and then analyzing how behavior over the cross-section of firms varies over the business cycle.

Steady state innovation and survival rates. Figure 1 highlights the extent of firm heterogeneity in our model. In particular, it shows firm profits, innovation rates, survival probabilities and the mass of firms as a function of (relative) firm-level productivity. In addition, it also explicitly depicts the values for young, medium and old incumbent firms.

Due to innovation, the support of the equilibrium distribution is moving to the right. As a result, old incumbents are on average less productive than medium age incumbents, which are less productive on average than young incumbents. This can be observed in Figure 1d. Moreover, since old firms are more efficient in production, conditional on their relative productivity $\widehat{q}$, they make larger profits as shown in Figure 1c. This has the direct implication that they have a higher survival probability and more incentives to undertake R&D as shown in Figures 1a and 1b, respectively.

As can be seen from the figure, our framework displays significant firm heterogeneity in terms of profits, with those of the most productive firms being around three times higher than those of the least. These, in turn, imply large differences in survival rates with young incumbent firms with low productivity having a 25 percent chance of shutting down, while old incumbents with high productivity facing less than a 5 percent chance of exit. That said, the mass of firms in these extremes is relatively low since unproductive firms get selected-out and highly productive firms are simply rare, as we can seen in Figure (1d). Importantly, innovation rates increase with productivity — we will return to this point when discussing

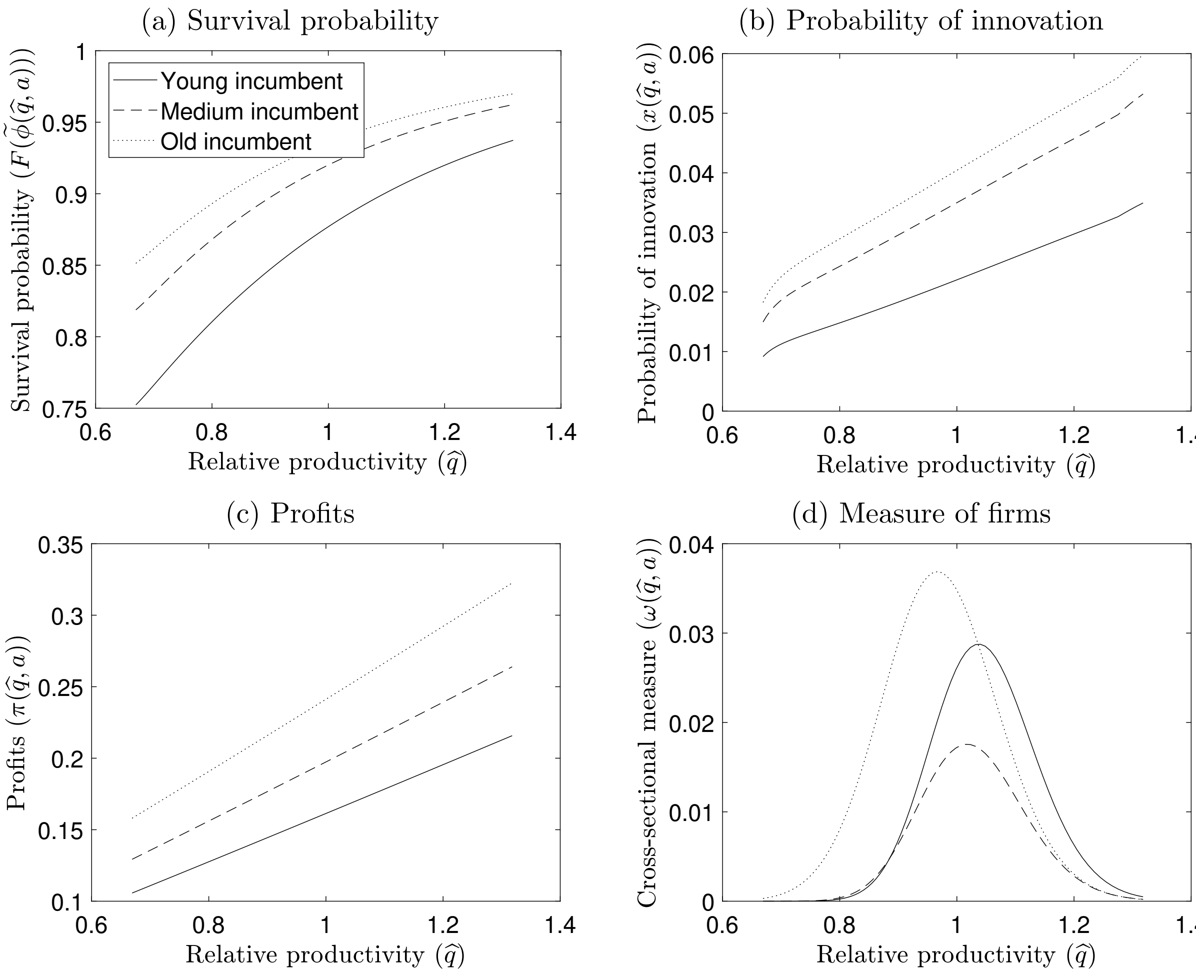

Figure 1: Incumbent heterogeneity in the steady state

aggregate dynamics. Appendix C depicts more properties of the incumbent distribution (Figure 5) and the entrant distribution over prototypes (Figure 6).

Cyclicality of innovation rates. Having shown that innovation rates differ across firms in the stationary steady state, we now turn to analyzing their cyclical behavior. In particular, we study the effect of a one standard deviation TFP shock on the impact responses, (i.e. the first period of their impulse response functions), of firm R&D spending over the cross-section of profits and productivities.[12]

Figures 2a and 2b show these responses of incumbent and entrant firms in the model in percentage deviations from steady state. There are two important angles to consider in

[12]We show how innovation cyclicaliity varies across firms differing in profits and profitability for completeness. Although, as shown in figure 1d more productive firms and also more profitable, so cyclicality varies similarly across the two firm characteristics.

this regard. The first is that, within an age grouping, these impact responses are declining monotonically with profits and productivity. The second is that the response profiles of firms are moving downwards as they move through their lifecycle and up the age groupings. Conditional on their type (profits/productivity), firms become less pro-cyclical as they age. One can think of older firms as being more profitable in our parameterization, given that they experience efficiency gains ψ_a that become larger as they mature.

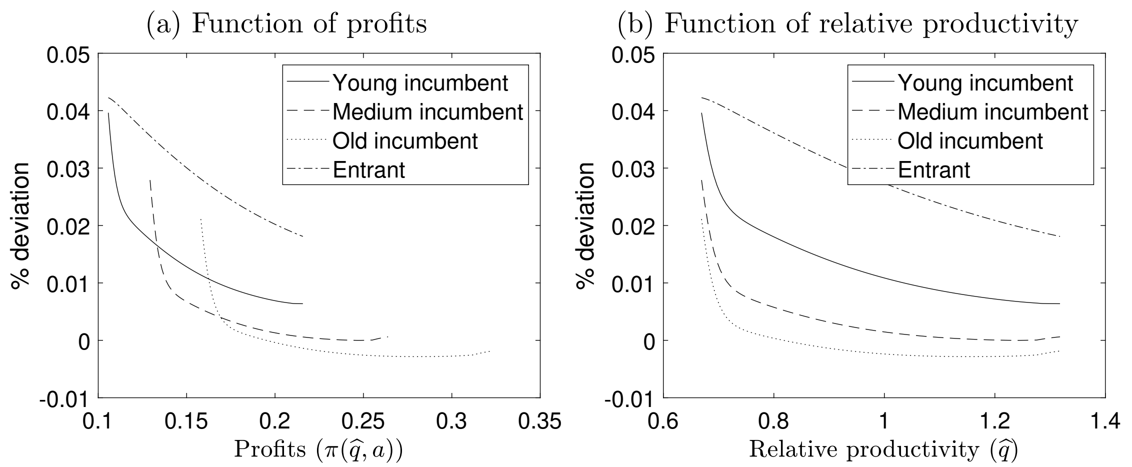

Figure 2: Impact responses of firm-level R&D spending from positive TFP shock. All responses are presented as percentage deviations from steady state (prior to multiplication by 100).

An intuitive explanation for this finding is that more productive firms are sufficiently far from the exit cutoff so that swings in their cash flows are less likely to threaten their future existence and incentive to grow. We will expand on this intuition in our analysis of the key mechanisms at play in the model, but before that we show that the model's theoretical prediction can be found in the data.

Empirical validation of cross-sectional innovation cyclicality. We now test our key prediction in the data. In particular, we take firm-level information from Compustat and industry-level output data from the Bureau of Economic Analysis (BEA). We construct a balanced-panel of firms over the period 1987–2019, with data at a yearly frequency.[13] Firms are placed into four groups based-on their profitability, relative to others in their industry; rankings are based on firms' real pretax income.[14] Given that firms can move around in the

[13]Given the requirement that the panel be balanced, we classify industries at the 2 digit level. We remove observations for firms in industries with fewer than ten firms reported over the sample period.

[14]We use the direct measure of profitability in Compustat, variable pi, and BEA inflation data. As a robustness, we also rank firms based-on profitability ratios: Compustat variables pi/sale. The inferences

profitability distribution over time, we construct groupings based on a preliminary regression of the form

$$v_{ijt} = \kappa_{1i} + \kappa_{2j} + \kappa_{3t} + e_{ijt} \tag{28}$$

where v_{ijt} is firm i's (in industry j) real profits at time t, κ_{1i} is a firm i fixed effect, κ_{2j} is a fixed effect for industry j to which firm i belongs, κ_{3t} is a time fixed effect and e_{ijt} is a residual. Object κ_{1i} gives an estimate of a firm's average lifetime profitability; we place firms into quartiles based on their estimates of κ_{1i}, relative to other firms in their industry. The main regression is then of the form

$$r_{ijt} = \alpha_0 + \alpha_1 y_{jt} + \sum_{k=2}^{4} \alpha_2^k g_{ik} + \sum_{k=2}^{4} \alpha_3^k g_{ik} y_{jt} + \epsilon_{ijt} \tag{29}$$

where r_{ijt} is firm i's (in sector j) growth in real R&D expenditures at time t, y_{jt} is their industry's gross real output growth and g_{ik} is an indicator for the firm being in quartile $k \in \{2, 3, 4\}$ for κ_{1i}. We consider various combinations of time and industry fixed effects.[15] Objects $\alpha_0, \alpha_1, \{\alpha_2^k\}_{k=2}^4, \{\alpha_3^k\}_{k=2}^4$ are parameters to be estimated; those of interest are α_1 and $\{\alpha_3^k\}_{k=2}^4$. Parameter α_1 gives the cyclicality of firms in the lowest grouping for profitability, while $\{\alpha_3^k\}_{k=2}^4$ are interaction terms of indicators for being in higher groupings with industry output growth. Regression results are given in Table 3.

Firms overall tend to be procyclical, with the magnitude of the relationship decreasing with the profitability quartile. One can infer the average cyclicality of a firm in quartile $k \in \{2, 3, 4\}$ by taking the sum $\alpha_1 + \alpha_3^k$. For all specifications, there are statistically significant differences across the four groupings. The least profitable firms are strongly procyclical (α_1), while the most profitable are close to acyclical in all specifications ($\alpha_1 + \alpha_3^4$). This regularity that more profitable firms are less pro-cyclical validates the key prediction of our model shown previously in Figure 2a. Having instilled confidence in the model's capacity to replicate salient features of the data on innovation heterogeneity, we now seek to explore the mechanisms that generate these predictions.

Economic mechanism: the role of selection. To shed more light on the role of selection in shaping the innovation cyclicality of firms, we first need to understand what drives R&D cyclicality in our economy. In our model, as in many endogenous growth models, innovation

are similar; we defer these results to Appendix D. This appendix also gives more details regarding the data cleaning process.

[15] We also consider firm-level fixed effects; the results are robust to their inclusion.

Table 3: Empirical validation regression results

Coefficient				
Real output growth (α_1)	2.283***	2.093***	2.282***	2.012***
	(0.280)	(0.311)	(0.282)	(0.297)
2^{nd} quartile interaction (α_3^2)	-1.670***	-1.667**	-1.667**	-1.666**
	(0.452)	(0.451)	(0.452)	(0.452)
3^{rd} quartile interaction (α_3^3)	-1.710***	-1.708***	-1.720**	-1.719**
	(0.376)	(0.376)	(0.371)	(0.370)
4^{th} quartile interaction (α_3^4)	-1.733***	-1.737**	-1.733**	-1.734**
	(0.364)	(0.359)	(0.364)	(0.356)
N	7,413	7,413	7,413	7,413
Time fixed effects (Y/N)	N	Y	N	Y
Industry fixed effects (Y/N)	N	N	Y	Y

Note: quartile groupings are done using κ_{1i} from regression (28). Table notation corresponds to regression equation (29). Standard errors are in parentheses. Superscripts *, ** and *** denote significance at the 10%, 5% and 1% confidence levels, respectively. N denotes number of observations. Dependent variable is firm-level real R&D growth.

generates intertemporal knowledge spillovers. When a firm innovates, it creates new ideas that future firms can use as a foundation for their own innovations. This positive externality results in firms underinvesting in innovation from a societal perspective. Additionally, it introduces a present bias in innovation: an innovating firm, aware that knowledge spillovers will allow other firms to eventually replace it, places less emphasis on long-term benefits compared to short-term gains. Consequently, entrepreneurs become shortsighted and highly responsive to short-term profit fluctuations. Barlevy (2007) elegantly illustrates this source of innovation cyclicality within a standard Schumpeterian growth model. Since innovation costs, such as wages, are procyclical, innovation itself becomes procyclical only if profit fluctuations exceed those of the costs. Barlevy (2007) demonstrates that high fixed costs result in significant profit fluctuations and identifies a threshold level of fixed operating costs that makes innovation procyclical.

Our model differs from the standard Schumpeterian growth model along three key dimensions. First, both incumbents and entrants innovate.[16] Secondly, firms must incur a sunk entry cost before engaging in innovation. Thirdly, significant firm heterogeneity exists, with firms varying in their equilibrium innovation rates and facing selection. All three factors play crucial roles in shaping the cyclical patterns of innovation at both aggregate and firm levels. Incumbent firms can secure future benefits by continuously innovating,

[16] In the standard Schumpeterian model only entrants innovate, due to the Arrow replacement effect (Aghion and Howitt, 1992; Grossman and Helpman, 1991) .

allowing them to survive and capitalize on investments made during downturns when conditions improve. Additionally, the cyclical nature of new entrants can potentially offset that of incumbents through a general equilibrium effect: procyclical entry increases the cost of innovation, thereby discouraging incumbents from boosting their innovation efforts during economic booms. Finally different firms could respond differently to shocks.

To illustrate the key mechanisms driving heterogeneity in innovation cyclicality in a simple and transparent way we consider versions of our model with varying magnitudes of the fixed cost.[17] In particular, we start with a low value of the distribution's mean, $\mu_{\phi 0}$ and increase the parameter to higher values $\mu_{\phi 1} > \mu_{\phi 0}$ and $\mu_{\phi 2} > \mu_{\phi 1}$. Figure 3a gives the analogue of Figure 2b for these three different values of μ_{ϕ}. Figure 3b is a version of 3a zooming on a shorter productivity interval that makes some results more transparent. Figure 3c gives the survival probabilities over the same cross-section of relative productivites for each configuration.

The first two values of μ_{ϕ} are insufficient to induce any degree of endogenous exit in the model: the survival probability is one, across the entire set of relative productivities. Nonetheless, a progressively higher cost burden for firms squeezes their level of profits. This translates into a progressively larger response of firms to the business cycle shock. The responses remain uniform for all firms, but rise monotonically, as shown clearly in figure 3b.

This aligns with the key finding in Barlevy (2007), which states that an increase in fixed operating costs makes firms' responses to a TFP shock more procyclical. Our analysis demonstrates that this holds true even in our more complex model, where innovation by incumbent firms leads to less shortsighted R&D decisions. In our model, though, firm heterogeneity adds a new margin of innovation cyclicality operating via selection. Intuitively, we can express the Bellman equation in (15) as

$$\widetilde{v}(\widehat{q}', a', Z') = F(\widetilde{\phi}(\widehat{q}', a', Z')) \left[v(\widehat{q}', a', Z') - \mathbb{E}[\phi | \phi < \widetilde{\phi}(\widehat{q}', a', Z')] \right] \tag{30}$$

which suggests that firms' continuation value is the product of the probability of survival $F(\widetilde{\phi}(\widehat{q}', a', Z'))$ and the next period value conditional on surviving. In the economic environment of Barlevy (2007), in which firms are symmetric and fixed costs are deterministic, the probability of survival along the business cycle is one for all firms and the continuation value driving firms' innovation decision is driven by the cyclicality of R&D costs and benefits (profits). Larger fixed production costs lead to stronger profit cyclicality and stronger R&D

[17] To better highlight the mechanism, we take a simplified version of the model with a single age grouping $n_A = 1$. More details on the parameterization can be seen from Appendix E, where we also use this simplified calibration for additional exercises.

cyclicality. We find a similar result in our richer model when the fixed cost is low enough to allow all firms to survive.

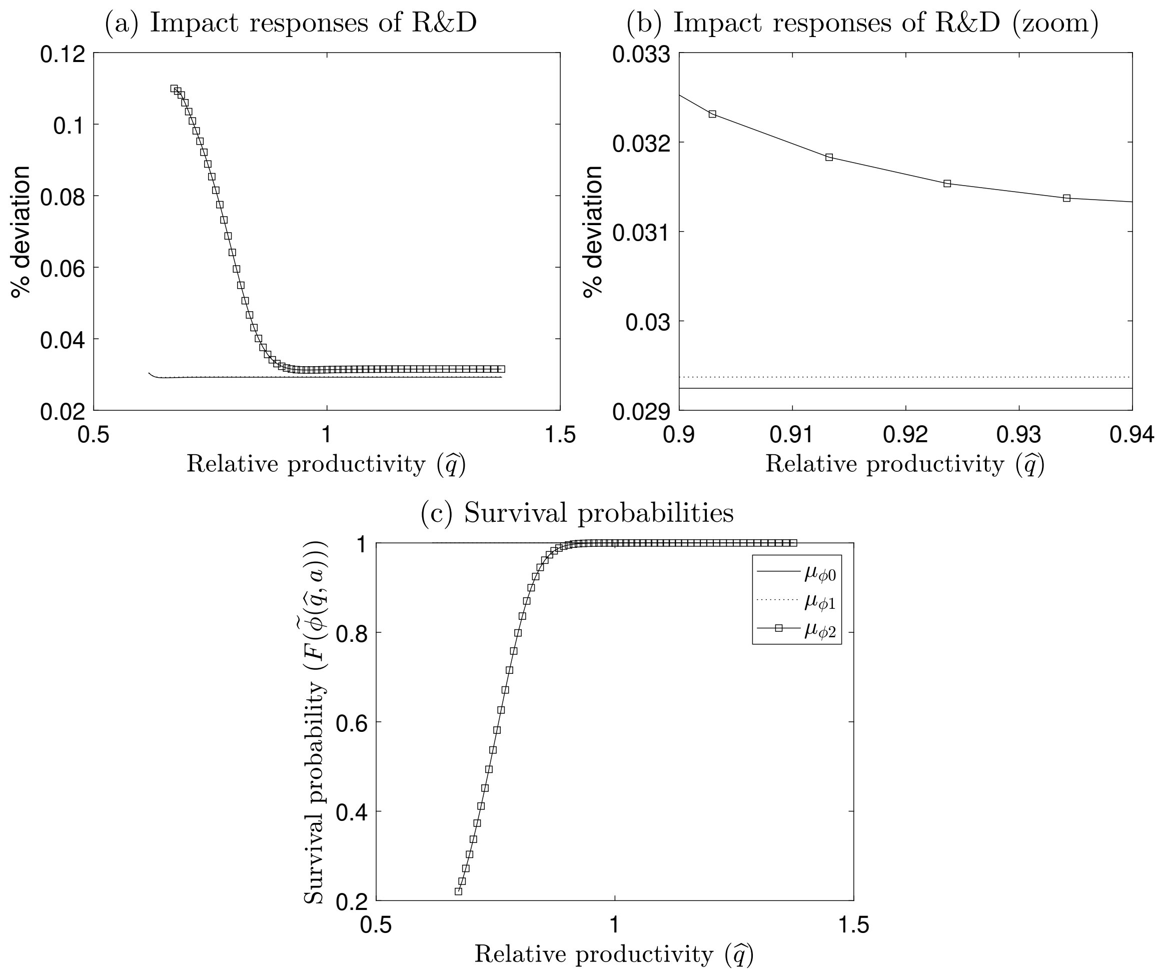

Figure 3: Fixed cost and selection

Eventually when the fixed cost mean crosses a threshold, endogenous exit is triggered for firms with relative productivity around a little below 1, as shown in figure 3c. This then leads to non-uniform responses of R&D on impact to the TFP shock. The response profile in figures 3a and 3b inherits a non-linear shape from the survival probability profile, where firms with low relative productivity are the most pro-cyclical, similarly to what we saw in figure 2a. This exercise illustrates in a transparent way the role of selection in shaping the differential responses of firms' innovation to business cycles. The survival probability is now one only for the most productive firms and affects the continuation value in (30), the more so the less productive the firms are. The risk of exit is higher for less productive firms which therefore respond more aggressively to the fluctuations in profits brought about by cyclical shocks.

Taking stock. The model generates significant heterogeneity in innovation rates and in innovation cyclicality across firms. While heterogeneity in firms' innovation rates is well-known, the differences in cyclicality is a novel prediction of our model. We provide new empirical evidence to support this theoretical prediction. Selection is the key driver of the stronger cyclicality of innovation for less productive firms predicted by the model.

4.2 Aggregate Growth

Having described the cyclical properties of our model, we now turn to its growth implications. Once again, we strive to quantify the impact of our main new elements – firm heterogeneity and firm selection. We do so by considering various counterfactual exercises in which we gradually shut-down our novel features.

Counterfactuals: heterogeneity in innovation rates and firm selection. As Figure 1b shows, there is substantial variation in innovation rates both within and across age groupings. For instance, the highest innovation probability is around 6 times larger than the smallest. To identify the effect of this feature, we conduct a counterfactual, referred to as CF_S, where we impose two common innovation levels across firms — one for incumbents and another for entrants instead of using the first order conditions (19) and (21). We calibrate these two common levels to hold the growth rate and entrant contribution to growth the same as in the baseline.[18] Along the transition, we then impose a common fluctuation across firms of each respective type, equal to the baseline's innovation fluctuation of an old incumbent and an entrant with $\widehat{q} = 1$.

Next, to isolate the role of entry and exit, we conduct a second counterfactual referred to as CF_{SMF}, where we also hold the entry mass M and exit probabilities $F(\widetilde{\phi})$ constant, on top of CF_S. That is all margins of innovation heterogeneity - levels, fluctuations, entry and exit - are all fixed in CF_{SMF}.

Results. Figure 4a presents the growth rate in percentage points, after a positive 1 standard deviation TFP shock, for the baseline as well as CF_S and CF_{SMF}. Notice substantial differences in the growth trajectories across the scenarios. More layers of heterogeneity typically lead to larger propagation of the shock. All three scenarios spike between 5–10 years,

[18] Note that, although untargeted, output volatility along the transition is roughly the same as in the baseline. More details relating to the computation of the counterfactual steady state are deferred to Step 12 of Appendix B.1.

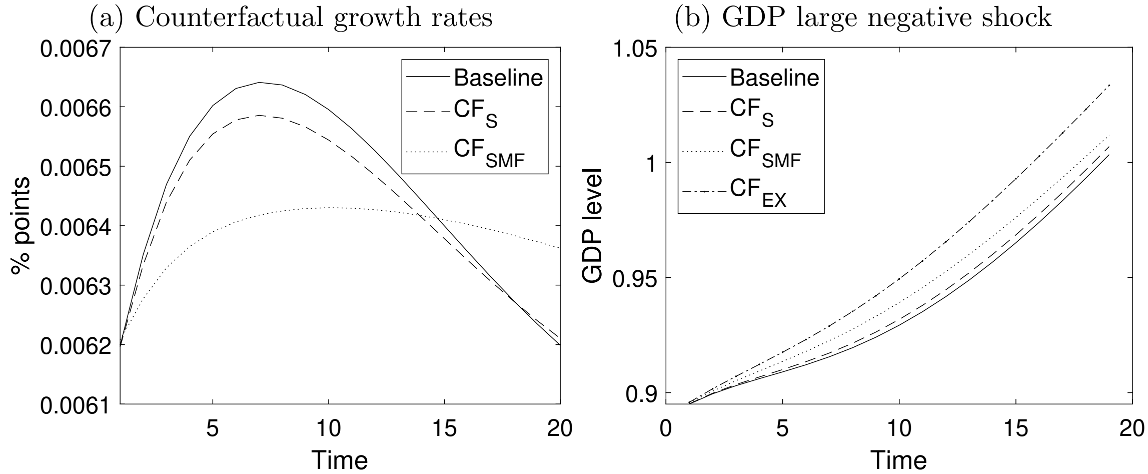

Figure 4: Counterfactual results. Percentages are prior to multiplication by 100.

with the baseline realizing a maximum of 0.665%. Removing heterogeneity in R&D levels lowers the peak to 0.655%, while also removing entry/exit drops it further to 0.640%.

Insights regarding these differences follow from re-allocations amongst the firm cross-section. Within each age grouping, a rise in the mass of the least productive firms ensues on impact and the periods immediately subsequent to the shock, coupled with a contraction of those of the most productive. Similarly, we see a rise in the mass of young firms, at the expense of medium and old. These re-allocations move in favor of firms that exhibit higher innovation rate cyclicality, as shown in Figure 2a. The culmination of this is an amplification of cyclicality in the growth rate in the baseline, relative to CF_S.

The intuition is relatively similar when inferring the effect of entry and exit. The measure of entrants surges on impact in the baseline, shifting mass towards young firms. We remove this effect when moving from CF_S to CF_{SMF}, abolishing this movement of resources from incumbents to entrants. Since entrants are the most pro-cyclical in their innovative responses to the shock (again see Figure 2a), this yields a weaker short-run growth effect. Hence, our model predicts procyclical entry and, symmetrically for a negative shock, countercyclical exit of less productive and younger firms, which could potentially weaken the propagation of technology shocks (as in e.g. Caballero and Hammour, 1994; Ottaviano, 2012). But, these firms are also the most procyclical in the innovation efforts which produces the opposite effect. In our framework and parametrization, differences in innovation cyclicality are the dominant force leading to a positive overall impact of firm heterogeneity on the propagation of technology shocks.

As a final simulation exercise, we implement a large negative shock to the economy and trace-out the time path followed by GDP; the results are shown in Figure 4b. In these

Table 4: Differences in average productivity

	1 year	10 Years	20 Years
Baseline difference from trend	-0.09	-2.13	-2.91
Heterogeneity effect	-0.04	-1.05	-0.73
Decomposition			
S heterogeneity	23.27	25.53	46.44
MF heterogeneity	76.73	74.47	53.56

Note: simulations with average productivity to 12 percent negative TFP shock at time 0. First line shows percentage deviation of baseline $Q(t)$ from trend growth (CF_{EX}). Simulation starts with $Q(0) = 1$ and rolls average productivity forwards using $Q(t) = Q(t-1)\exp(G(t-1))$ for $t \geq 1$. Second line measures differences between baseline and CF_{SMF} $Q(t)$ levels at cutoff time period, as a percentage of the CF_{SMF} $Q(t)$ level. Decomposition of S heterogeneity then gives differences in $Q(t)$ predictions of CF_S-CF_{SMF} as fraction of overall difference. MF difference is residual. All numbers are percentages after multiplication by 100.

simulations, we administer a one time negative shock to TFP of 12%. We choose this shock size so that the baseline simulation realizes an impact drop in real GDP of around 10%, in order to mimic the magnitude of losses experienced in the onset of the COVID-19 pandemic. In this figure, we depict the time paths followed by the baseline, CF_S and CF_{SMF}, in addition to another counterfactual, labeled as CF_{EX}, in which we have trend growth. That is, the latter counterfactual sees a variable time path given the shock's effect on stationarized output, while growth simply continues at the steady state rate.

Figure 4b has the simulation starting at 1, where all scenarios then see a drop to around 0.9 at the shock's impact. This figure highlights that endogenous growth and our novel features of heterogeneity have a quantitatively significant impact on key aggregates of the economy. The overall effect of endogenous growth (baseline v.s. CF_{EX}) on GDP is around 4% after 20 years, that is, GDP is still 4% below what it would have been under trend (steady-state) growth. Firm heterogeneity accounts for roughly one percentage point of this difference (baseline v.s. CF_{SMF}). Consequently, entry and exit, along with the heterogeneous innovation response of firms, explain approximately $1/4^{th}$ of the shock's persistence after 20 years.

We then perform a decomposition of the trajectory of the average productivity variable $Q(t)$ after the large negative shock in Table 4. Average productivity is still 2% lower than with trend growth ten years after the shock and almost 3% lower 20 years after. This indicates that the impact of a large TFP shock on productivity in our endogenous growth environment is very persistent. Heterogeneity gives average productivity that is around 1% lower after the ten year mark, while dropping slightly to 0.73% after 20 years. Both elements of heterogeneity play a meaningful role in shaping this difference, with innovation heterogeneity explaining around 1/4 up to around 10 years, but rising in prominence to around 1/2 subsequently.

4.3 Costs of Business Cycles Revisited

As a final step in our analysis, we revisit the classic question of costs of business cycles. Following Lucas (1987) and in line with the existing literature, we adopt the standard methodology of estimating the welfare costs of business cycles. In particular, we compare our baseline economy to a hypothetical one in which business cycles are eliminated — our deterministic steady state. To quantify the welfare impact, we compute the fraction of per-period consumption, ζ, which agents in our model would be willing to give up in order to remain in the deterministic steady state as opposed to the stochastic steady state with business cycles. This welfare loss measure, ζ, is computed from:

$$\sum_t \beta^t U(C_t) = \sum_t \beta^t U(C_t^*(1-\zeta)),$$

where an asterisk indicates the deterministic steady state.

In other words, we compute the net present value of utility in our baseline model with and without aggregate shocks. The value ζ that makes households indifferent between one and the other is the welfare cost of business cycles. In these calculations, we consider several alternative specifications of our model. First, the baseline in which growth is endogenous, it reacts to business cycles and firm heterogeneity plays an important role. Second, CF_{SMF}, where recall we shut-down heterogeneity in innovation as well as holding entry and exit fixed along the cycle. Finally, we also consider a model with exogenous growth. To generate these numbers, we simulate 25 time series of 200 Monte Carlo TFP shocks, solving for firm choices and the implied cross-section, then averaging across the welfare numbers that follow.

Our findings indicate that business cycle fluctuations lead to significant economic losses, with households willing to sacrifice approximately 3% of their annual consumption to avoid these fluctuations. Without heterogeneity, that is abstracting from entry and exit and heterogeneity in innovation, the cost of business cycle is instead 2.6%, which is 15% lower than in our baseline model.

These costs are several orders of magnitude larger than the 0.1% found by Lucas (1987, 2003). Barlevy (2004) show that endogenous growth can lead to even larger costs of fluctuations than those we find, with 7-8% of annual consumption. But his exercise keeps average investment fixed in the fluctuating economy, so that the baseline model and the counterfactual economy have the same initial consumption. While in our model, as in previous models that have calculated the cost of business cycle in endogenous growth models, fluctuations affect investment in innovation and initial consumption. With this calculation, it is notoriously harder to obtain large welfare costs of fluctuations.[19] Not only is the overall cost of fluctuations quite large, considering the conservative methodology used to compute it, more importantly, the new channel of welfare losses that we identify of firm heterogeneity, provides a non-negligible contribution.

5 Conclusion

Our study underscores the importance of integrating growth and business cycle dynamics, particularly in light of recent crises such as the late 2000s global financial meltdown and the COVID-19 pandemic. By developing a model that combines endogenous growth with business cycle fluctuations and heterogeneous firms, we reveal critical insights into the role of innovation and firm heterogeneity. Our findings indicate that innovation is inherently procyclical, especially for smaller, less productive, and less profitable firms, a novel prediction supported by empirical evidence from Compustat data. This heightened procyclicality among these firms is driven by their present-bias in innovation and greater sensitivity to short-term profit fluctuations due to exit risks during economic downturns.

Our counterfactual analyses demonstrate that firm heterogeneity significantly amplifies the transmission of total factor productivity (TFP) shocks. When firm innovation levels and cyclicality are equalized, or when endogenous entry and exit are removed, the model shows reduced volatility and diminished shock amplification. In contrast, the baseline model, which incorporates firm heterogeneity, indicates that a 10% GDP drop, akin to the COVID-19 shock, results in GDP remaining 4% below long-run trend twenty years later, with firm heterogeneity accounting for a quarter of this persistent gap. Procyclical entry and countercyclical exit of less productive firms, while potentially mitigating technology shocks, also contribute to greater business cycle amplification through their procyclical innovation behavior.

Our results challenge the traditional view posited by Lucas (1987), which suggested that the costs of economic fluctuations are minimal. We find that the inclusion of firm

[19] If eliminating fluctuations increases investment in growth-enhancing activities, the economy will indeed grow faster. But since such investment takes resources away from consumption, average initial consumption will fall, and the benefits of eliminating fluctuations will be smaller (e.g Epaulard and Pommeret, 2003).

heterogeneity raises the estimated costs of business cycles to approximately 3% of annual consumption, significantly higher than Lucas's estimate of less than 0.1%. Excluding firm heterogeneity reduces this cost to 2.6%, highlighting the substantial role that firm diversity plays in exacerbating the welfare costs of business cycles.

In conclusion, our study highlights the critical need to consider firm heterogeneity in models of economic growth and business cycles. This approach provides a more nuanced understanding of the propagation mechanisms of macroeconomic shocks and highlights the significant welfare implications of cyclical fluctuations, driven by the diverse responses of firms.

One fruitful area for future research is the role of policy interventions in mitigating the adverse effects of business cycle fluctuations on long-term growth (see e.g. Benigno and Fornaro, 2018; Fornaro and Wolf, 2023). Investigating the effectiveness of fiscal and monetary policies in supporting innovation and firm survival during economic downturns could inform more nuanced policy designs.

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, reallocation, and growth," *American Economic Review*, 2018, *108* (11), 3450–3491.

_ , _ , Nicholas Bloom, and William Kerr, "Innovation, Reallocation and Growth," *American Economic Review*, 2018, *108* (11), 3450–3491.

Aghion, Philippe and Gilles Saint-Paul, "On the Virtue of Bad Times: An Analysis of the Interaction between Productivity Growth and Economic Fluctuations," *Macroeconomic Dynamics*, 1998, *2*, 322–344.

_ and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

_ , David Hemous, and Enisse Kharroubi, "Cyclical Fiscal Policy, Credit Constraints, and Industry Growth," *Journal of Monetary Economics*, 2014, *62*, 41–58.

_ , Marios Angeletos, Abhijit Banerjee, and Kalina Manova, "Volatility and Growth: Credit Constraints and Productivity-Enhancing Investment," *Journal of Monetary Economics*, 2010, *57*, 24–67.

— , **Nicolas Berman, Laurent Eymard, Philippe Askenazy, and Gilbert Cette**, "Credit Constraints and the Cyclicality of R&D Investment: Evidence from France," *Journal of the European Economic Association*, 2012, *10*, 1001–1024.

— , **Philippe Bacchetta, Romain Rancierre, and Kennet Rogoff**, "Exchange Rate Volatility and Productivity Growth: The Role of Financial Development," *Journal of Monetary Economics*, 2009, *56*, 494–513.

Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez, "Endogenous Technology Adoption And R&D as Sources of Business Cycle Persistence," *American Economic Journal: Macroeconomics*, forthcoming.

Ates, Sina T. and Felipe Saffie, "Fewer but Better: Sudden Stops, Firm Entry, and Financial Selection," Federal Reserve Bank Working Paper 2018.

Ball, Lawrence, "Long-Term Damage from the Great Recession in OECD Countries," *NBER Working Paper*, 2014, (20185).

Barlevy, Gadi, "The Cost of Business Cycles under Endogenous Growth," *American Economic Review*, 2004, *94* (4), 964–990.

— , "On the Cyclicality of Research and Development," *American Economic Review*, 2007, *97* (4), 1131–1164.

Benigno, Gianluca and Luca Fornaro, "Stagnation Traps," *Review of Economic Studies*, 2018, *85*, 1425–1470.

Bianchi, Francesco, Howard Kung, and Gonzalo Morales, "Growth, slowdowns, and recoveries," *Journal of Monetary Economics*, 2019.

Bilbiie, Florin O, Fabio Ghironi, and Marc J Melitz, "Endogenous entry, product variety, and business cycles," *Journal of Political Economy*, 2012, *120* (2), 304–345.

Bloom, Nicholas, Rachel Griffith, and John van Reenen, "Do R&D Tax Credits Work?," *Journal of Public Economics*, 2002, *85* (1), 1–31.

Boucekkine, Raouf, Marc Germain, and Omar Licandro, "Replacement echoes in the vintage capital growth model," *Journal of economic theory*, 1997, *74* (2), 333–348.

— , **Omar Licandro, Luis A Puch, and Fernando Del Rio**, "Vintage capital and the dynamics of the AK model," *Journal of economic theory*, 2005, *120* (1), 39–72.

Caballero, Ricardo J. and Mohamad L. Hammour, "The Cleansing Effect of Recessions," *The American Economic Review*, 1994, *84* (5), 1350–1368.

Cerra, Valerie, Antonio Fatas, and Sweta Saxena, "Hysteresis and Business Cycles," *Journal of Economic Literature*, 2023, *2023*, 181–215.

Comin, Diego and Mark Gertler, "Medium-Term Business Cycles," *American Economic Review*, 2006, *96* (3), 532–551.

Cozzi, Guido, Beatrice Pataracchia, Philipp Pfeiffer, and Marco Ratto, "How Much Keynes and How Much Schumpeter?," *European Economic Review*, 2021, *133*, 103–660.

Epaulard, Anne and Aude Pommeret, "Recursive utility, endogenous growth, and the welfare cost of volatility," *Review of Economic Dynamics*, 2003, *6* (3), 672–684.

Fatas, Antonio, "Do Business Cycles Cast Long Shadows? Short-run Persistence and Economic Growth," *Journal of Economic Growth*, 2000, *5*, 147–162.

Fornaro, Luca and Martin Wolf, "The scars of supply shocks: Implications for monetary policy," *Journal of Monetary Economics*, 2023, *140*, S18–S36.

Foster, Lucia, John Haltiwanger, and Cornell J Krizan, "Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s," *The Review of Economics and Statistics*, 2006, *88* (4), 748–758.

Ghironi, Fabio and Marc J. Melitz, "International Trade and Macroeconomic Dynamics with Heterogeneous Firms*," *The Quarterly Journal of Economics*, 2005, *120* (3), 865–915.

Griliches, Zvi, "Patent Statistics as Economic Indicators," *Journal of Economic Literature*, 1990, *28*.

Grossman, Gene and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 1991, *58* (3), 43–61.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," 2001. NBER Working Paper 8498.

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

Krebs, Tom, "Growth and welfare effects of business cycles in economies with idiosyncratic human capital risk," *Review of Economic Dynamics*, 2003, *6* (4), 846–868.

_ , "Job Displacement Risk and the Cost of Business Cycles," *American Economic Review*, 2007, *97* (3), 664–686.

Krusell, Per and Anthony A Smith Jr, "Income and wealth heterogeneity in the macroeconomy," *Journal of political Economy*, 1998, *106* (5), 867–896.

_ , Toshihiko Mukoyama, Ayşegül Şahin, and Anthony A. Smith, "Revisiting the welfare effects of eliminating business cycles," *Review of Economic Dynamics*, 2009, *12* (3), 393–404.

Lee, Yoonsoo and Toshihiko Mukoyama, "Entry and Exit of Manufacturing Plants over the Business Cycle," *European Economic Review*, 2015, *77*, 20–27.

Lucas, Robert, "Macroeconomic Priorities," *American Economic Review*, 2003, *93* (1), 1–14.

Lucas, Robert E. Jr., *Models of Business Cycle*, New York: Basil Blackwell, 1987.

Moscarini, Giuseppe and Fabien Postel-Vinay, "The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment," *American Economic Review*, 2012, *102* (6), 2509–2539.

Nuno, Galo, "Optimal Research and Development and the Cost of Business Cycles," *Journal of Economic Growth*, 2011, *16*, 257–283.

Olivier, Eugenio Cerruti Summers Lawrence Blanchard, "Inflation and Activity—Two Explorations and Their Monetary Policy Implications," *NBER Working Paper*, 2015, (21726).

Ottaviano, Gianmarco I. P., "Firm Heterogeneity, Endogenous Entry, and the Business Cycle," *NBER International Seminar on Macroeconomics*, 2012, *8* (1), 57–86.

Ouyang, Min, "On the Cyclically ofR&D," *Review of Economics and Statistics*, 2011, *93*, 542–533.

Pancost, N Aaron and Chen Yeh, "Decomposing Aggregate Productivity," Technical Report 2022.

Schmitz, Tom, "Endogenous Growth, Firm Heterogeneity and the Long-run Impact of Financial Crises," *European Economic Review*, 2021, *132* (2).

Sedláček, Petr, "Creative destruction and uncertainty," *Journal of the European Economic Association*, 2020, *18* (4), 1814–1843.

Sedláček, Petr, "The Lost Generation of Firms and Aggregate Labor Market Dynamics," *Journal of Monetary Economics*, 2020, *111*, 16–31.

_ **and Vincent Sterk**, "The Growth Potential of Startups over the Business Cycle," *American Economic Review*, 2017, *107* (10), 3182–3210.

Spencer, Adam Hal, "Policy effects of international taxation on firm dynamics and capital structure," *The Review of Economic Studies*, 2022, *89* (4), 2149–2200.

Stiglitz, Joseph, "Endogenous Growth and Cycles," in Yuici Shionoya and Mark Perlman, eds., *Innovation in Technology, Industries, and Institutions*, University of Michigan Press, 1994, pp. 126–156.

Tallarini, Thomas, "Risk-Sensitive Real Business Cycles," *Journal of Monetary Economics*, 2000, *45* (3), 507–32.

Vinci, Francesca and Omar Licandro, "Switching-Track after the Great Recession," 2020. CFCM Working paper 02/2020.

A Equilibrium Definition

In this appendix we detail the formal expressions for the equilibrium growth rate and full sequence equilibrium definition.

A.1 Growth Rate

We define the growth rate of the average quality level in the economy, drawing on the cross-sectional measure law of motion (25)–(26). We denote the growth rate as $G' = Q'/Q - 1$. From the definition of Q in (7), the net growth rate is defined in variety-level notation as

$$(1 + G')^{\eta-1} = \frac{1}{\Omega'} \int_{j\in\Omega'} \left(\frac{q'}{Q}\right)^{\eta-1} dj. \tag{31}$$

Leveraging the definition of the cross-section (25)–(26) in (31), we expresses the growth rate in terms of state variables as

$$
\begin{aligned}
\Omega'(1+G')^{\eta-1} = & \sum_{a\in\{1,2,\ldots,n_A-1\}} \zeta_a(1-\delta_a) \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}(1+\lambda)/(1+G'),a+1,Z')} [\widehat{q}(1+\lambda)]^{\eta-1}\, x\left(\widehat{q},a,Z\right) \omega\left(\widehat{q},a,Z\right) \mathrm{d}F(\phi) d\widehat{q} \\
& + \sum_{a\in\{1,2,\ldots,n_A-1\}} \zeta_a(1-\delta_a) \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}/(1+G'),a+1,Z')} \widehat{q}^{\eta-1} \left[1-x\left(\widehat{q},a,Z\right)\right] \omega\left(\widehat{q},a,Z\right) \mathrm{d}F(\phi) d\widehat{q} \\
& + \sum_{a\in\{1,2,\ldots,n_A-1\}} (1-\zeta_a)(1-\delta_a) \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}(1+\lambda)/(1+G'),a,Z')} [\widehat{q}(1+\lambda)]^{\eta-1}\, x\left(\widehat{q},a,Z\right) \omega\left(\widehat{q},a,Z\right) \mathrm{d}F(\phi) d\widehat{q} \\
& + \sum_{a\in\{1,2,\ldots,n_A-1\}} (1-\zeta_a)(1-\delta_a) \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}/(1+G'),a,Z')} \widehat{q}^{\eta-1} \left[1-x\left(\widehat{q},a,Z\right)\right] \omega\left(\widehat{q},a,Z\right) \mathrm{d}F(\phi) d\widehat{q} \\
& + (1-\delta_{n_A}) \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}(1+\lambda)/(1+G'),n_a,Z')} [\widehat{q}(1+\lambda)]^{\eta-1}\, x\left(\widehat{q},n_a,Z\right) \omega\left(\widehat{q},n_a,Z\right) \mathrm{d}F(\phi) d\widehat{q} \\
& + (1-\delta_{n_A}) \int_{\widehat{q}} \int^{\widetilde{\phi}(\widehat{q}/(1+G'),n_a,Z')} \widehat{q}^{\eta-1} \left[1-x\left(\widehat{q},n_a,Z\right)\right] \omega\left(\widehat{q},n_a,Z\right) \mathrm{d}F(\phi) d\widehat{q} \\
& + M \int_{\widehat{q}_e} \int^{\widetilde{\phi}(\widehat{q}_e(1+\lambda)/(1+G'),1,Z')} \widehat{q}_e(1+\lambda) x_e\left(\widehat{q}_e,Z\right) \omega_e\left(\widehat{q}_e\right) \mathrm{d}F(\phi) d\widehat{q}_e. \qquad (32)
\end{aligned}
$$

Expression (32) considers all possible relative quality levels at time t in the first of each of the double integrals. Firms then move up or down from their starting points, based-on their choices of x and x_e and the corresponding measures of firms. Note that we need to consider age groups below the top category separately from those below, as well as entrants, who arrive in the lowest age group.

A.2 Stochastic Stationary Equilibrium Definition

A stochastic stationary equilibrium is an infinite sequence of aggregate objects

$$\{\widehat{C}_t, A_{t+1}, \widehat{W}_t, R_t, \widehat{Y}_t, \Lambda_t, M_t, \Omega_t, Z_t, G_t\}_{t=0}^{\infty},$$

firm-level incumbent variables

$$\{\{k(\widehat{q}_t,a_t,Z_t), v(\widehat{q}_t,a_t,Z_t), \widetilde{v}(\widehat{q}_t,a_t,Z_t), \widetilde{\phi}(\widehat{q}_t,a_t,Z_t), x(\widehat{q}_t,a_t,Z_t), s(\widehat{q}_t,a_t,Z_t), \omega(\widehat{q}_t,a_t,Z_t)\}_{(\widehat{q}_t,a_t,Z_t)}\}_{t=0}^{\infty}$$

and entrant variables

$$\{\{v_e(\widehat{q}_{e,t}, Z_t), x_e(\widehat{q}_{e,t}, Z_t), s_e(\widehat{q}_{e,t}, Z_t)\}_{(\widehat{q}_{e,t}, Z_t)}\}_{t=0}^{\infty}$$

such that the following conditions hold

1. The households make their optimal consumption-savings choice with $\widehat{W}_t$ and R_t taken as given, to yield Euler equation (10),
2. The stochastic discount factor Λ_t is defined from the household's solution (10),
3. The aggregate TFP process Z_t evolves as in (5),
4. The value functions $v_t(\widehat{q}_t, a_t, Z_t)$ in (13) and $\widetilde{v}_t(\widehat{q}_t, a_t, Z_t)$ in (15) solve the firm's optimisation problem, with optimal static control vector

   $$k(\widehat{q}_t, Z_t) = \{p(\widehat{q}_t, Z_t), n(\widehat{q}_t, Z_t), \widehat{y}(\widehat{q}_t, Z_t)\}$$

   of prices, employment and de-trended output, exit threshold $\widetilde{\phi}(\widehat{q}_t, Z_t)$ in (18) and innovation probability $x(\widehat{q}_t, Z_t)$ in (19),
5. The value function $v_e(\widehat{q}_{e,t}, Z_t)$ in (20) solves the entrant's problem, with optimal innovation probability $x_e(\widehat{q}_{e,t}, Z_t)$ in (21),
6. The measure of entrants is determined by (22).
7. The cross-sectional measure of firms evolves over time as in (25)–(26),
8. Markets clear consistent with firm decisions, the endogenous cross-sectional measure and aggregate conditions given by (23) and (24).
9. The endogenous growth rate is given in (32).

B Computational Algorithms

In this appendix, we detail the solution of the deterministic steady state and transition after a shock, in turn.

B.1 Deterministic Steady State

The following describes the computation of the calibrated steady state. Given it is deterministic, we omit time subscripts. Note the distinguishing characteristic here is that aggregate TFP is fixed $Z = 1$.

1. Create a grid for relative productivity levels $\widehat{q}$ around the average Q.

2. Fix certain aggregates: this reduces the iterative complexity of the problem. Specifically, we fix the de-trended wage $\widehat{W} = 1$ as well as de-trended aggregate output $\widehat{Y} = 1$. To make these aggregates consistent with our equilibrium definition in Appendix A.2, we treat the sunk cost of entry c_e and the overall firm mass Ω as free parameters. Recall also that the aggregate price index is the numéraire in our model, as such we also set $P = 1$.

3. Solve for the static choices in vector $k(\widehat{q}, a, 1)$ of intermediate goods producing incumbents consistent with Step 2, as well as their model-implied expressions in (16), (17) and the de-trended version of (11).

4. Solve the incumbent intermediate producing firms' Bellman equation (13) using value function iteration. This uses the static controls from Step 3 as an input. This step yields their policy functions for R&D $x(\widehat{q}, a, 1), s(\widehat{q}, a, 1)$, cutoff fixed cost $\widetilde{\phi}(\widehat{q}, a, 1)$, expected fixed costs, probability of exit and value functions $v(\widehat{q}, a, 1)$, $\widetilde{v}(\widehat{q}, a, 1)$. Note that we extrapolate the value function beyond the top grid point, so that firms getting close to the maximum $\widehat{q}$ still have an incentive to innovate, as in the model.

5. Solve the entrant's problem using the value function $\widetilde{v}(\widehat{q}, 1, 1)$ found in Step 4 as an input and the expressions given in (20) and (21). This yields their optimal R&D choices $x_e(\widehat{q}, 1)$, $s_e(\widehat{q}, 1)$ as well as the entrant value conditional on prototype $v_e(\widehat{q}, 1)$.

6. Find the sunk cost of entry consistent with the free entry condition, the object $v_e(\widehat{q}, 1)$ found in Step 4 and the assumption that $\widehat{W} = 1$ from Step 2, $c_e = \mathbb{E}[v_e(\widehat{q}, 1)]$.

7. Find the stationary endogenous cross-section of firms across age and relative productivity. Do this by firstly assuming a unit measure of firms (distribution). Then writing equations (25)–(26) in matrix notation

$$\omega_t = \Gamma_t \omega_{t-1} + M_t \omega_{e,t} \tag{33}$$

where ω_t a vector of the distribution of firms across the state space, Γ_t is an endogenous Markov transition matrix, consistent with the equilibrium objects found in Steps 3–5, M_t is still the measure of entrants and $\omega_{e,t}$ is the initial distribution of prototype productivities. We can then find the stationary invariant distribution through inverting (33) as $\widetilde{w} = \widetilde{M}(I-\Gamma)^{-1}\omega_e$ where I is the identity matrix, $\widetilde{\omega}$ is the stationary distribution and $\widetilde{M}$ is the measure of entrants that normalizes the overall firm measure sum to unity.

8. Find the variable averages, across all intermediate producing firms, that are consistent with the distribution from Step 7 and optimal controls from Steps 3–5.

9. Find the measure of entrants that is consistent with the assumption that $\hat{Y} = 1$ from Step 1 using the average de-trended output from Step 8 and stationary distribution from Step 7, leveraging the linearity of the latter. This also then implies the cross-sectional measure of incumbents across the state space.

10. Compute the aggregate variables consistent with Step 9. Find the aggregate labor demand from intermediate goods firms, N^D. Set the inelastic labor supply of the household equal to this object $N = N^D$. Treat this as a parameter for the transition simulations.

11. Find the aggregate growth rate as described in Appendix A.1 and contribution from entrants.

12. Find endogenous objects relating to the steady state of the counterfactuals CF_S and CF_{SMF}. Impose a common $\bar{x}$ and $\bar{x}_e$ for incumbents and entrants, respectively. Use a solver to find these objects to match the same growth rate and entrant contribution to growth found in Step 11. The counterfactual steady state uses the same static $k(\widehat{q}, a, 1)$ and exit choices $\widetilde{\phi}(\widehat{q}, a, 1)$ of incumbents that are found in Steps 3–4 of the baseline steady state computations. We then impose that $x(\widehat{q}, a, 1) = \bar{x}$ and $x_e(\widehat{q}, 1) = \bar{x}_e \; \forall \widehat{q}, a$. Then taking these policy functions, we compute the invariant distribution as in Step 7. When finding the mass of firms, we impose that the measure of entrants in the counterfactual steady state is the same as in the baseline; the overall measure is then implied.

B.2 Transition After TFP Shock

To solve for the dynamic equilibrium, we build on Sedláček (2020), but extend the methodology to allow for endogenous growth. In particular, we use first-order perturbation around the stationary steady state (including the steady state *life-cycle* patterns of firms). The first-order approximated solutions, denoted by hats, have the following form:

$$\begin{aligned}\widehat{h}_{t+1} =& \overline{h} + \Theta(\widehat{h}_t - \overline{h}),\\ \widehat{d}_{t+1} =& \overline{d} + \Pi(\widehat{h}_t - \overline{h}).\end{aligned}$$

where Θ and Π are matrices containing the coefficients obtained from the approximation, $\widehat{h}_t$ contains deviations of state variables and $\widehat{d}_t$ is a vector containing deviations of non-predetermined variables. The perturbation procedure is standard and carried out in one step.

An advantage of perturbation methods is that the computational speed is relatively high and many state variables can be handled.[20] An important prerequisite for perturbations to be accurate, however, is that deviations from the steady-state are not too large.

For firm dynamics models like the one in this paper it may seem problematic because differences in employment levels across firms may be very large. The solution method adopted here, however, overcomes this problem since the steady state we perturb around contains the entire life-cycle profiles of firms. These growth paths, captured by the constants in the above equations, are themselves non-linear functions of firms's productivity.

Hence, the fact that most newborn firms starts off much below their eventual sizes does not involve large accuracy losses since the same is true for the steady-state sizes of newborn firms. Similarly, the fact that the equilibrium features various firm types with very different optimal sizes does not reduce accuracy since we perturb around the growth path for each individual firm type.

C Additional Figures on Steady State Heterogeneity

Figure 5 depicts additional properties of the incumbent firm distribution, while Figure 6 displays properties of the entrant distribution, conditional on their prototype draw.

[20] Computational time is significantly lower when compared with global methods for dealing with aggregate shocks such as Krusell and Smith (1998) and after MIT policy shocks such as Spencer (2022).

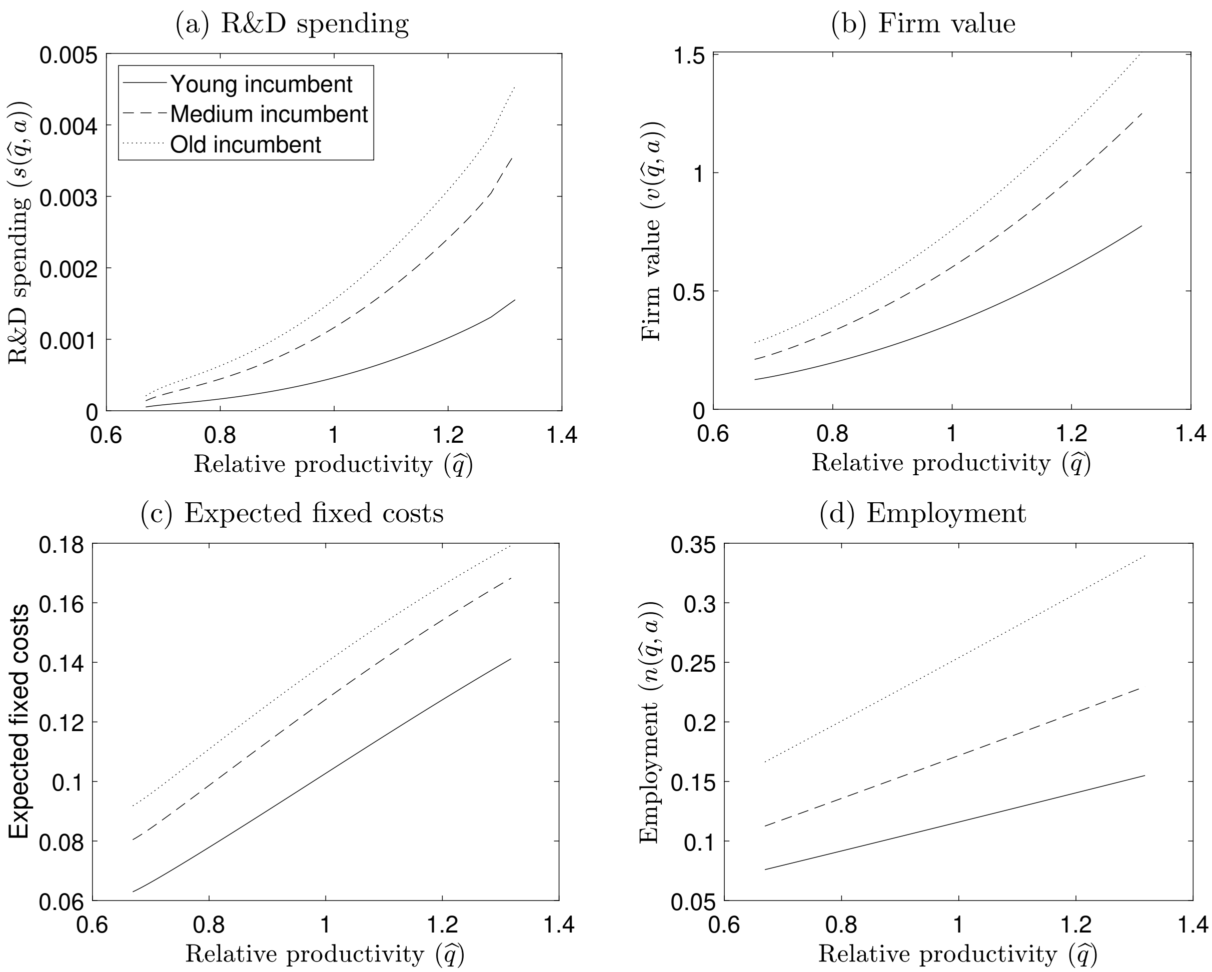

Figure 5: Additional figures on incumbent heterogeneity in the steady state

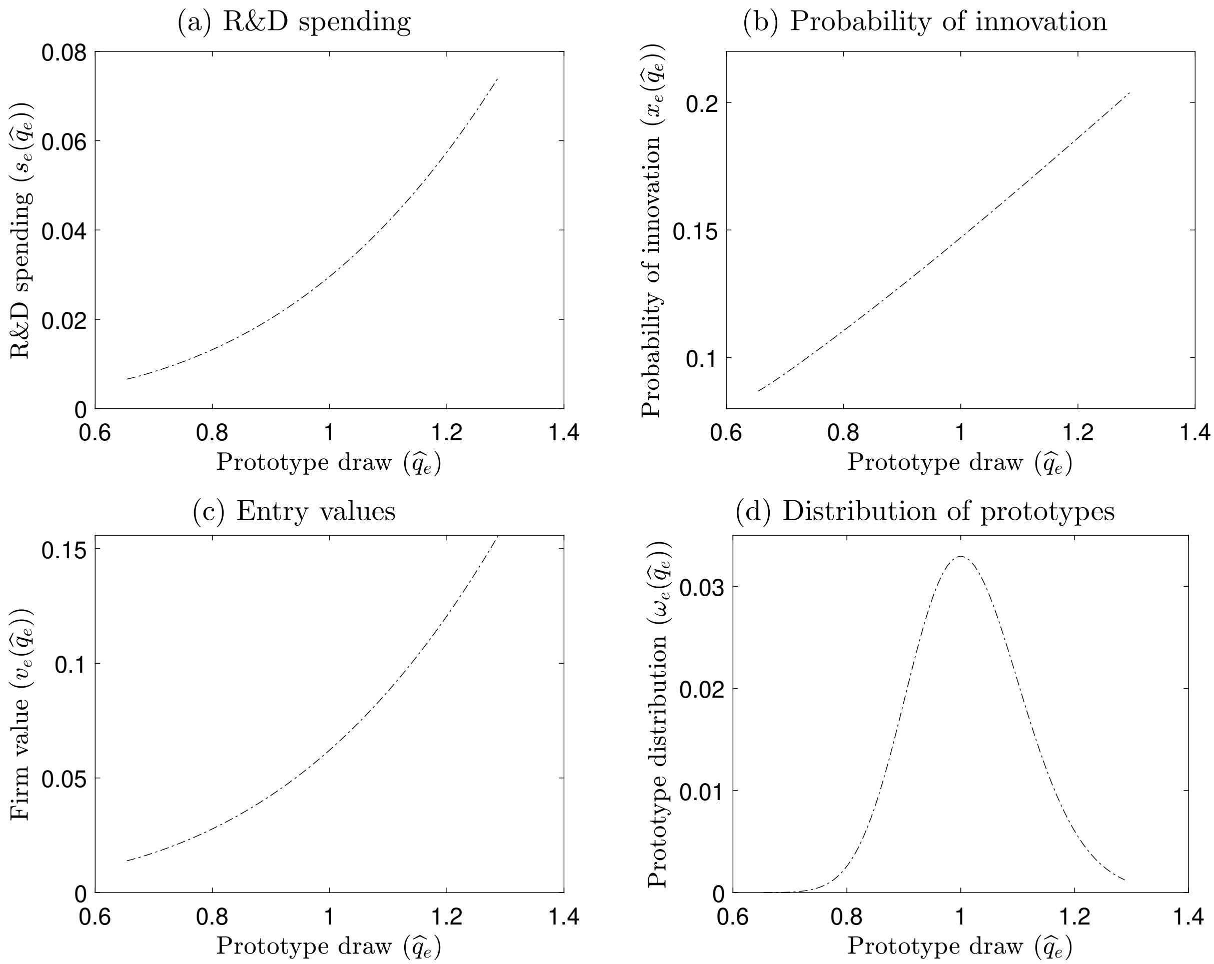

Figure 6: Entrant heterogeneity in the steady state

D Data Appendix

Table 5 presents results for the empirical regressions with ranking based on profitability ratios, rather than real profits. We follow the below procedure to generate the regression coefficients given in Table 3 in the text, as well as Table 5.

1. Load CPI data from FRED series CPIAUCSL. Take the average CPI across all months within a given year to find a yearly measure. Use 1987 as the base year for deflating.
2. Load Compustat Annual data. Extract calendar years and merge with the CPI data from Step 1.
3. Only study U.S. firms: drop those without a listed foreign incorporation code (fic) or headquarters code (loc). Only keep firms whose entries of these two indicators are the USA.
4. Drop observations with old industrial codes (indfmt == "FS").
5. Drop observations that are in non-USD currencies (curcd != "USD").
6. Drop observations outside the range of years considered.
7. Drop firms that have double-reporting of data for a given fiscal year.
8. Drop if firms have double-reporting over calendar years.
9. Drop firms that change their reporting month over the sample.
10. Define the real growth rate of a firm using its reported income statement item xrd, in conjunction with CPI data created in Step 1.
11. Extract 2 digit NAICS codes from firms. Drop firms whose industry changes over the sample.
12. Download and merge with industry-level gross output data from the BEA. Deflate using CPI information from Step 1.
13. Impose a balanced panel for firms for all years under consideration.
14. Drop all firms in an industry if there are fewer than 10 reported in total.
15. Generate firm lifetime profitability estimate using fixed effects as in (28) with Stata function xtreg. Output the firm fixed effect. Pretax income (pi) divided by sales (sale) is the variable used for sorting (v_{ijt}).

Table 5: Empirical validation regression results: profitability ratio robustness

Coefficient				
Real output growth (α_1)	2.460***	2.212***	2.484***	2.200***
	(0.271)	(0.295)	(0.258)	(0.275)
2^{nd} quartile interaction (α_3^2)	-1.850**	-1.844**	-1.881**	-1.876**
	(0.434)	(0.436)	(0.412)	(0.415)
3^{rd} quartile interaction (α_3^3)	-1.986***	-1.978**	-2.017***	-2.009***
	(0.366)	(0.368)	(0.345)	(0.345)
4^{th} quartile interaction (α_3^4)	-2.031***	-2.032**	-2.065***	-2.064***
	(0.377)	(0.373)	(0.353)	(0.349)
N	7,413	7,413	7,413	7,413
Time fixed effects (Y/N)	N	Y	N	Y
Industry fixed effects (Y/N)	N	N	Y	Y

Note: rankings instead based on profitability ratios with Compustat variables pi/sale. Quartile groupings are done using κ_{1i} from regression (28). Table notation corresponds to regression equation (29). Standard errors are in parentheses. Superscripts *, ** and *** denote significance at the 10%, 5% and 1% confidence levels, respectively. N denotes number of observations. Dependent variable is firm-level real R&D growth.

16. Generate groupings (quartiles) of the fixed effects estimated in Step 15 by industry.

17. Give firms labels, based on how their fixed effect estimated in Step 15 places relative to the cutoffs estimated in Step 16.

18. Run the regressions presented in Table 3.

E Single Age Grouping Calibration and the Role of Entry

The parameterization in this appendix takes a simplified calibration with only one single age grouping ($n_A = 1$). The parameter values used are given in Table 6 with the corresponding moments in Table 7. We calibrate this version of the model to contain purely exogenous exit, for the purposes of the exposition, leveraging small variance σ_ϕ and relatively low fixed cost mean μ_ϕ, such that there is no endogenous exit. We therefore calibrate 7 parameters to 7 data targets. All other relevant parameters remain the same as in the baseline.

Now using this simplified calibration, we explore the role of the elasticity of entry χ on the impact responses of incumbents' R&D. We start with the assumption of free entry, represented through χ of ∞. We then gradually reduce this parameter, until the point of

Table 6: Parameter values

Parameter		Value	Source/Target
Discount factor	β	0.96	Annual interest rate 4%
Coefficient of relative risk aversion	σ	1.00	Logarithmic preferences
TFP persistence	ρ_Z	0.78	Sedláček and Sterk (2017)
Innovation function curvature	ψ	2.00	Hall et al. (2001); Bloom et al. (2002)
Entrant prototype mean	μ_e	0.00	Normalisation/baseline
Elasticity of entry	χ	15.0	Baseline
TFP volatility	σ_Z	0.02	Real GDP volatility
Elasticity of substitution	η	4.43	Profits/GDP ratio
Innovation step size	λ	0.05	Growth rate
Incumbent innovation productivity	γ	2.07	Aggregate R&D to GDP
Entrant prototype variance	σ_e	0.07	Mean prod. entrants/incumbents
Entrant innovation productivity	γ_e	4.30	Entrant contribution to growth
Exogenous death rate	δ	0.11	Exit rate

a zero elasticity, at which point entry is non-responsive to changes in value after a shock. Figure 7a shows the cross-section of responses to a one standard deviation TFP shock, while Figure 7b gives the trajectory of the measure of entrants.

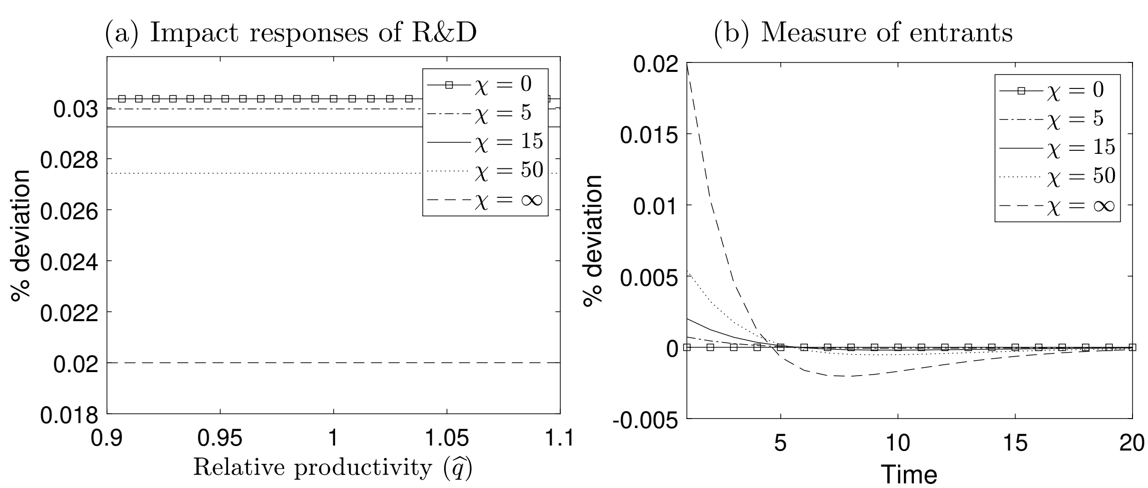

Figure 7: Entry elasticity

As the elasticity becomes lower, the responses of R&D become more pro-cyclical. The intuition being that incumbents seek to expand their innovative efforts during a boom, but entry is a competing process in terms of utilization of resources. When this entry process is restricted, there is more capacity for incumbent firms to increase their R&D. As such, changing this parameter translates the R&D responses of firms up or down.

Table 7: Targeted moments

Moment	Data	Model	Source
Volatility of real GDP	0.02	0.02	BEA
Overall Exit rate	0.11	0.11	BDS
Profits/GDP ratio	0.12	0.14	BEA
Productivity growth	0.82	0.82	BLS
Aggregate R&D to GDP	0.04	0.08	BEA
Mean prod. entrants/incumbents	1.02	1.01	Foster et al. (2006)
Entrant contribution to growth	0.39	0.39	Pancost and Yeh (2022)

Note: all moments are targeted. All numbers are prior to multiplication by 100, except for the average productivity growth: this is presented as a percentage after x100.

A more detailed intuition of the mechanism at work is the following. As we saw, entry is highly procyclical and leads to general equilibrium effects that crowd-out incumbents' response to shocks. The entry process requires labor through the sunk entry cost and units of final goods used in the innovation process. If entry is very elastic and therefore very procyclical, wages will also be very procyclical thereby taming the procyclical response of incumbents who experience stronger increases in costs during booms and stronger reductions in recessions. Moreover, if entry is strongly procyclical more final goods are demanded for entrants' innovation during booms, thereby squeezing resources for innovation by incumbents. This leads to higher marginal utility of consumption and so lowers investment by incumbents via the stochastic discount factor.