Investment, Innovation, and Financial Frictions*

Pablo Ottonello
University of Michigan and NBER
pottonel@umich.edu

Thomas Winberry
Wharton and NBER
twinb@wharton.upenn.edu

May 31, 2023

Abstract

We develop a framework in which firms grow by accumulating capital to scale up existing ideas or innovating to create new ideas. Empirically, we find that capital investment rates decline with size and age while innovation rates increase with size and age. Our model matches this pattern due to financial frictions; constrained firms grow by accumulating capital because of its high marginal product and collateral value, while unconstrained firms grow by innovating because they have exhausted the returns to capital. Financial frictions delay the point at which firms start innovating and therefore lower long-run growth. To the extent that ideas are non-rival, the equilibrium allocation features too little innovation and too much investment. We find that subsidizing innovation approximately corrects this allocation. Cutting taxes on investment can also raise innovation in the long run, but at the cost of lowering innovation in the short run.

*We thank Andy Abel, Marco Bassetto, Gideon Bornstein, Jeff Buckley, Jeremy Greenwood, Erzo Luttmer, Simon Mongey, Guillermo Ordonez, Victor Rios-Rull, Pierre-Daniel Sarte, and seminar participants at the Jackson Hole Macro Conference 2023, the Penn-Wharton Brownbag, the Minneapolis Fed, the University of Wisconsin, and the Richmond Fed. We thank Samuel Stern and Erin Gibson for excellent research assistance.

1 Introduction

Our goal in this paper is to understand the role of financial frictions in driving aggregate economic growth. Of course, aggregate growth is the result of purposeful decisions of individual firms to grow themselves. We study two key margins along which firms grow: accumulating inputs, like capital, or innovating to create new ideas. While input accumulation is necessary to implement existing ideas, new ideas push out the technological frontier and are the only source of sustained growth in the long run. In this context, financial frictions will create a tradeoff in how firms allocate their available funds between these two margins of growth.

Our main contribution is to develop a framework to study the allocation of investment and innovation across heterogeneous firms subject to financial frictions. In our model, firms trade off the benefit of accumulating capital — its marginal product and collateral value — against the benefit of innovating — increasing the likelihood of a new idea. This tradeoff naturally generates a "pecking order" in which constrained firms primarily grow through investment but unconstrained firms grow primarily through innovation. We show that this pecking order is consistent with a new fact in the microdata: capital investment rates are declining in size and age, while innovation rates are increasing in size and age.

If the ideas produced by innovation are non-rival, the equilibrium allocation will be inefficient, opening the door to policy intervention. A constrained-efficient planner who internalizes the non-rivalry of ideas would raise innovation, but may raise or lower investment depending on the degree of financial constraints across firms and over time. While simple policies cannot exactly replicate this allocation, we find that a constant innovation subsidy can get very close. If that subsidy is not completely available, we also find that more generous tax deductions for investment—such as the full expensing enacted by the Tax Cuts and Jobs Act of 2017— successfully increase growth by raising the returns to innovation in the long run, but at the expense of lower innovation in the short run due to even higher investment.

Our model combines elements of the Hopenhayn (1992) framework, in which firm dynamics are determined given an exogenous process for productivity, and the endogenous growth framework, in which productivity is determined through innovation. We primarily focus on incumbent firms who face the tradeoff between investment and innovation; new entrants simply draw a new idea from the existing stock of ideas in the economy.[1] Incumbent firms must decide how much resources to spend on investment, which increases their capital stock, and innovation, which increases the probability of receiving a new idea and permanently increasing their productivity. There are four key differences between capital and ideas in our model: (i) idea arrival is risky but capital accumulation is not, (ii) capital is collateralizable but ideas are not, (iii) capital is sellable but ideas are not, and (iv) ideas are partially non-rival but capital is not. Constrained firms favor investment over innovation because its high marginal product, collateral value, and lower risk.

Our model's pecking order of firm growth is consistent with two new facts in firm-level data from Compustat and Orbis: capital investment rates are declining in size and age, while innovation rates — measured using R&D expenditures or successful patent applications — are increasing in size and age.[2] Financial frictions are necessary for our model to match these patterns; without financial frictions, firms would immediately jump to their optimal scale, leaving investment and innovation counterfactually independent of size and age.

A key challenge in calibrating the model is disciplining the innovation technology, which governs the expected returns to innovation. While we can arguably measure the inputs into the innovation technology using R&D expenditures, the outputs (new ideas) are difficult to directly measure in the data. Instead, we use the occurrence of investment spikes to reveal the arrival of new ideas. Consistent with our model, R&D expenditures are a strong predictor of investment spikes in firm-level data from Compustat 1975-2018. We use the regression coefficient of investment spikes on R&D expenditures to pin down how the probability of success varies with innovation expenditures, and then use the average size of investment spikes pins down the size of successful innovations

Our calibrated model implies that financial frictions lower the long-run growth rate compared to a version of the model without financial frictions (in which small firms immediately

[1]While financial frictions may distort innovation by new entrants as well, they do not face the tradeoff between scaling up their existing ideas vs. creating new ideas that we study in this paper. Garcia-Macia, Hsieh and Klenow (2019) estimate that the majority of aggregate growth is driven by innovation among incumbents rather than new entrants.

[2]We compute these patterns using within-sector or within-firm variation in order to control for permanent technological differences across sectors/firms (which are absent from our model). If we instead look across sectors and firms, we find that innovation rates are increasing in size, consistent with Akcigit and Kerr (2018).

begin innovating and contributing to growth). Of course, financial frictions also dampen capital accumulation and create a misallocation of capital across firms, which lowers the level of aggregate productivity as in Buera, Kaboski and Shin (2011) or Midrigan and Xu (2014). However, over long horizons, the bigger costs of financial frictions come from a lower growth rate due to lower innovation. In addition, because financial frictions prevent small firms from innovating, the majority of innovation activity in our model is concentrated among the largest, most unconstrained firms in the economy.[3]

In order to understand the differences between the private and social returns to innovation, we study a constrained-efficient allocation chosen by a planner who is subject to the same financial constraints as individual firms.[4] Unlike individual firms, the planner internalizes how firm-level innovations increase aggregate productivity; financial frictions amplify this externality because higher aggregate productivity also alleviates financial constraints. We compute the transition path chosen by the planner following a similar approach to Lucas and Moll (2014).

While the planner always wants higher innovation, there is a meaningful tradeoff for investment: higher innovation requires lower investment for constrained firms but incentivizes higher investment for unconstrained firms (due to the complementarity between productivity and capital). In the early phase of the transition, the substitutability for constrained firms dominates and the planner lowers aggregate investment. Over time, the resulting growth builds up the distribution of net worth, and eventually the complementarity for unconstrained firms dominates in the sense that aggregate investment increases.

We use these insights to evaluate the growth effects of simple policies often implemented in practice. Unfortunately, these simple policies cannot exactly decentralize the planner's allocation because that allocation varies over time and across firms. However, we find that a innovation subsidy that is constant across firms and time gets very close to the planner's

[3]This finding implies that financial crises, modeled as an unexpected tightening of financial constraints, do not have an especially persistent impact on aggregate growth because they do not significantly reduce aggregate innovation.

[4]The constrained-efficient approach is common in models with incomplete markets because it does not endow the planner with the power to arbitrarily complete markets in a way that the private sector cannot. Conversely, this approach takes as given that the planner cannot impact whatever underlying frictions lead to these missing markets. To the extent that the planner can indeed alleviate those frictions, our results provide a lower bound for the true welfare gains from the optimal policy.

allocation in terms of aggregate variables. This occurs because constrained firms endogenously have a high marginal propensity to innovate out of the subsidy, which requires them to lower investment.

We also show that an investment tax cut can *partially* substitute for the innovation subsidy in the sense that the tax cut also succeeds at raising innovation and growth in the long run, but at the expense of lowering innovation in the short run. The tax cut raises long-run growth by raising the return to innovation among unconstrained firms and freeing up cash flows for constrained firms (by lowering after-tax expenditures on investment). However, the tax cut leads to a surge of higher investment initially raises the real interest rate, which temporarily lowers in the incentive to innovate. These results contrast with the neoclassical growth model, in which investment tax cuts have no effect on long-run growth due to the diminishing marginal product of capital.[5]

Related Literature As described above, our model weaves together elements of the Hopenhayn (1992) and endogenous growth frameworks. A key feature of Hopenhayn (1992) is decreasing returns to scale, which implies that firms have an optimal scale given their level of productivity. The following literature has studied how various frictions impede the ability of firms to reach this optimal scale, such as firing costs in Hopenhayn and Rogerson (1993), adjustment costs in Khan and Thomas (2008), selection upon entry in Clementi and Palazzo (2016), or — most closely related to our model — financial frictions in Khan and Thomas (2013). We contribute to this literature by incorporating innovation, which endogenizes the productivity process and therefore the distribution of optimal firm size.

On the endogenous growth side, our focus on firm dynamics is related to the creative destruction literature pioneered by Aghion and Howitt (1992) and Grossman and Helpman (1991), and more recently used in quantitative analyses by, e.g., Klette and Kortum (2004), Akcigit and Kerr (2018), or Acemoglu et al. (2018). Most of these models abstract from frictions to factor accumulation and, therefore, have no source of sluggish dynamics.[6] We contribute to this literature by incorporating capital accumulation and sluggish dynamics

[5] Howitt and Aghion (1998) also argue that subsidizing investment can stimulate innovation in a related model in which capital is an input in the R&D process.

[6] An important exception is Bilal et al. (2021), who study a version of the creative destruction models with labor in which search frictions generate sluggish dynamics.

induced by financial frictions.

Our empirical work regarding investment spikes is related to the large literature on lumpy investment (see, for example, Caballero and Engel (1999), Cooper and Haltiwanger (2006), or Khan and Thomas (2008)). We contribute to this literature by showing that investment spikes are systematically related to R&D expenditures, at least among the firms in our Compustat sample. This result leads us to reinterpret some investment spikes as coming from successful innovations, allowing us to use them to infer the innovation technology.

Finally, our paper relates to the literature on optimal policy in models with financial frictions. This literature focuses on how pecuniary externalities from borrowing behavior on interest rates can lead to suboptimal outcomes; see, for example, Geanakoplos and Polemarchakis (1986); Kehoe and Levine (1993); and Lorenzoni (2008). We contribute to this literature by studying a different externality: how innovation decisions alleviate other firms' financial constraints. In this respect, our results echo Itskhoki and Moll (2019), who study how wage suppression alleviates financial constraints in developing economies.

Road Map The rest of the paper is organized as follows. Section 2 presents evidence from the microdata that investment rates decline with size and age while innovation rates increase with size and age. Motivated by this evidence, Section 3 develops the model. Section 4 shows that firms in the model grow according to the pecking order described above, consistent with our motivating evidence. Section 5 calibrates the model and Section 6 shows how financial frictions affect growth and the firm-size distribution. Section 7 shows how the non-rivalry of ideas opens the door to policy intervention and evaluates the growth effects of innovation subsidies and investment tax cuts. Section 8 describes potential extensions of the framework and Section 9 concludes.

2 Motivating Evidence

This section describes the empirical allocation of investment and innovation among incumbent firms, which are the focus of our model.

Data Description We use annual firm-level data from Compustat and Orbis. Our baseline analysis uses data from Compustat, a panel of publicly listed U.S. firms that satisfies two key requirements for our analysis. First, it contains joint information on firms' investment, R&D expenditure, and financial positions, which allows us to measure the relationships of interest for our analysis. Second, it is a long panel, which allows us to absorb technological differences across firms — which are absent in our model – using fixed effects. The main disadvantage of Compustat is that it excludes private firms. To address this limitation, we complement the analysis with Orbis, which also features data from privately held firms.

Our main variables of interest are firms' investment rates, which we measure as the ratio of capital expenditures (x_{jt}) to lagged plant, property, and equipment (k_{tj}); and R&D-to-sales, which we measure as the ratio of research and development expense ($\widehat{i}_{jt} = A_t z_{jt} i_{jt}$) to the average of sales in the previous 5 years ($\tilde{y}_{jt} \equiv \frac{1}{5}\sum_{k=1}^{5} y_{jt-k}$). We complement this measure of innovation expenditure with data on patents from the United States Patent and Trademark Office collected by Kogan et al. (2017). Appendix C.1 provides the definitions of all variables used in our empirical analysis as well as details on our sample selection, which focuses on the period from 1975 to 2018.[7]

Results We document how investment and innovation rates vary across firms' size and age. Since our model abstracts from technological differences across firms, we focus our analysis within narrowly-defined sector by demeaning all variables at the 4-digit NAICS level. We measure age using the years since incorporation (available from Datastream) and size using log capital.

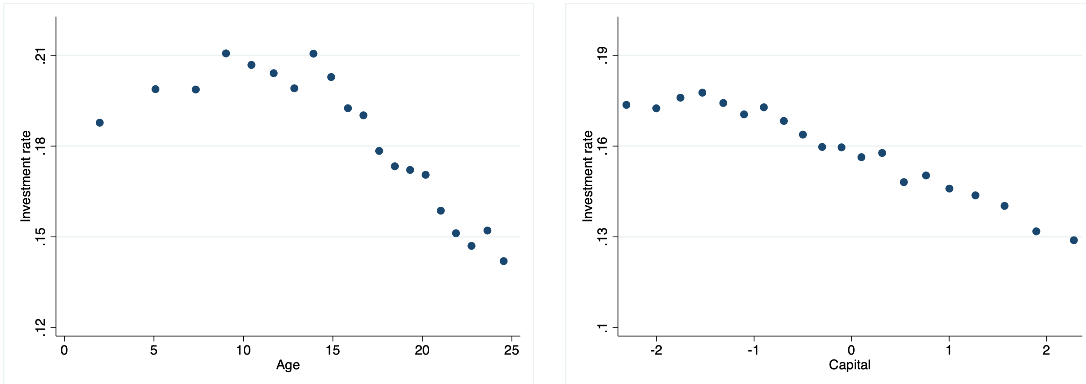
We document two key patterns in Figure 1. First, Panel (a) shows that capital investment rates are decreasing in size and age. For example, young firms' investment rates are nearly two thirds higher than old firms, while small firms' investment rates are around one third higher. This finding is consistent with the idea from our model that small and young firms face a high return to capital.

Our second key finding, in panels (b) and (c), is that innovation rates are increasing in size and age. We report the share of firms within each size/age group that report positive

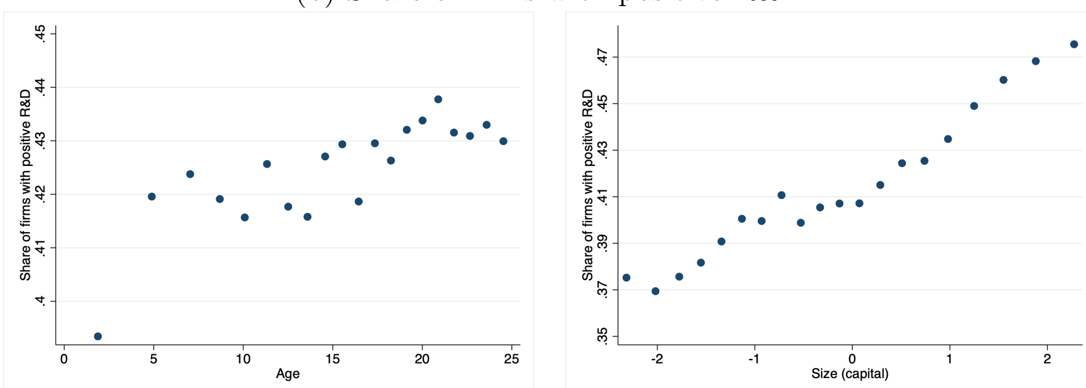
[7]Following Peters and Taylor (2017), we start our analysis in 1975, when the Federal Accounting Standards Board started requiring firms to report their R&D expenditure.

FIGURE 1: Investment Rates and Innovation by Firms' Size and Age

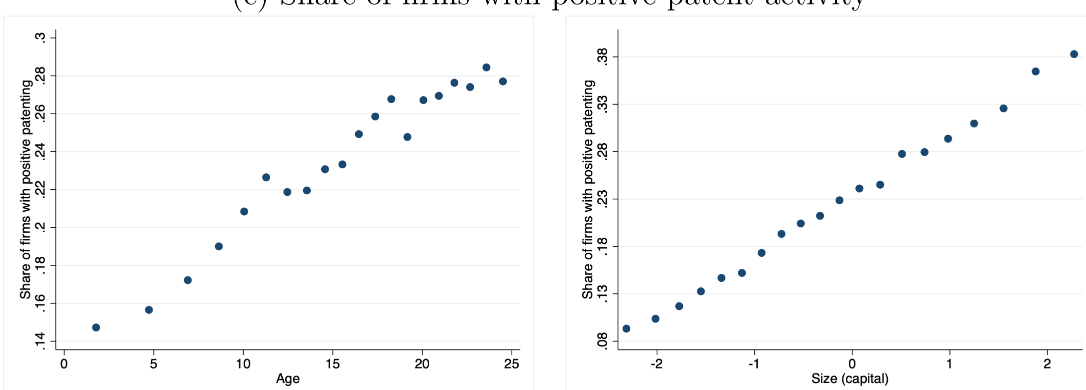
(a) Investment rates



(b) Share of firms with positive R&D



(c) Share of firms with positive patent activity



Notes: These figures report binned scatter plots of investment rates, the share of firms with positive R&D, and the share of firms with positive patenting by firms' size (measured by the log of real capital) and age of incorporation. All variables are demeaned at 4-digit-NAICS sector level. To construct the plots for investment rates, the share of firms with positive R&D rates, the share of firms with positive patenting, and age variables, we add the unconditional mean of each variable to sector-level demeaned variables. For plots involving age, we restrict observations to firm-years with ages between 0 and 25 years. For variable definitions and sample selection, see Appendix C.1.

R&D expenditures in panel (b) and who report patenting activity in Panel (c). Panel (b) shows that young firms are around 12% less likely to actively engage in R&D than old firms and that small firms are 22% less likely than large firms. Similarly, Panel (c) shows that young firms are about half as likely to patent in a given year as old firms, while small firms are only 20% as likely to patent as large firms.[8] These findings are consistent with the idea that innovation only becomes profitable once the firm has driven down its returns to capital.

Appendix Figure 10 shows that these patterns by size are even stronger if we use purely within-firm, as opposed to within-sector, variation (though at the cost of reducing the number of observations in order to have enough within-firm variation). In contrast, using all the variation *across* firms and sectors implies that R&D and patent rates instead *increase* with size, as in Akcigit and Kerr (2018).[9] Again, we prefer to focus on within-sector/firm dynamics because our model abstracts from technological differences across firms.

While our Compustat sample is selected relative to the economy as a whole, we think that it nevertheless provides a useful representation of the growth dynamics of the typical innovating firm in our model for two related reasons. First, Appendix Figure 12 shows similar patterns if we use Orbis data, which contains some privately held firms. Second, using Census microdata that includes many more of the smallest private firms in the economy, Acemoglu et al. (2018) find that the vast majority of innovation is done by the largest firms.

3 Model

We now develop our model of investment and innovation that is consistent with the evidence presented above. The model is set in discrete time and there is no aggregate uncertainty.

[8]Appendix Figure 1 shows that, conditional on pursuing positive R&D, old firms also tend to have larger R&D-to-sales firms than young firms with positive R&D, though R&D-to-sales ratios for firms with positive R&D do not exhibit a systematic pattern across the size distribution.

[9]To facilitate the comparison of our results with those in Akcigit and Kerr (2018), Appendix Figure 11 considers a data treatment following that in their work, reporting statistics using the "raw data" (i.e., without demeaning by sector or firm) and in the sample "continuously-innovative firms" (i.e., those that in the previous five years conducted some positive R&D expenditure or patenting). We confirm their findings that R&D-to-sales ratio is decreasing in firms' size and patents per employee are decreasing in firms' size hold in their sample applied to our data as well.

3.1 Environment

Before presenting the details of the model, we find it useful to briefly describe the types of firms we study. The majority of firms in the data are *non-ambitious* in the sense that they pursue little to no innovation and their optimal scale is very small, perhaps for non-pecuniary reasons (see Hurst and Pugsley (2011)). We omit these firms from our analysis and instead focus on *ambitious* firms, who eventually innovate and meaningfully contribute to economic growth. We conceptualize the lifecycle of these ambitious firms in two phases. In the first *pre-entry phase*, an entrepreneur uses their own time and skills to attempt to generate a new idea. Once a new idea materializes, the entrepreneur creates a firm and enters the *incumbent phase*. In this phase, the firm must decide how much to scale up its existing idea through investment vs. innovating to attempt to generate a new idea. Our model focuses on this second phase, which accounts for nearly all of the firm's life and, Garcia-Macia, Hsieh and Klenow (2019) estimate, accounts for the majority of growth in the economy.

New Entrants We model the pre-entry phase by simply assuming that there is a fixed flow π_d of new entrants each period who are endowed with zero debt and draw their initial levels of productivity and capital from some distribution $\Phi_t^0(z, k)$. In order to capture imitation by new entrants (as in, e.g., Luttmer (2007)), we assume this initial distribution of productivity is related to the distribution of incumbent firms in the economy; we will parameterize and calibrate this dependence in Section 5. Imitation is one sense in which ideas are non-rival and is an externality that incumbents do not internalize.

Incumbents Firms in the incumbent phase, indexed by j, produce output $y_{jt} = (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha}$ where A_t is aggregate productivity, z_{jt} is firm-specific productivity, and k_{jt} is the firm's capital stock (inherited from past investment decisions).[10] Decreasing returns to capital $\alpha < 1$ ensure there is an optimal scale of the firm for each level of productivity, as in Hopenhayn

[10] One can view this production function as one in which static inputs, like labor, have already been optimally chosen. In this case, the production parameter α should be interpreted as the elasticity of output to this combination of inputs. We are currently working on incorporating variable labor into the model, which implies that innovation generates a negative pecuniary externality through higher wages (see Section 8). Importantly, the growth rate of aggregate productivity in this extended model is the same as in the current model without labor.

(1992).[11] At the beginning of the period, a random subset of firms learn that they must exit the economy, in which case they produce, sell their undepreciated capital $(1-\delta)k_{jt}$, and pay back their existing debt. This exit shock occurs with probability π_d, the same as the flow of new entrants, which ensures the mass of firms in production is 1 at all times.[12]

Firms that will continue into the next period spend resources on two sources of growth, investment and innovation. Investment expenditures x_{jt} yield capital in the next period following the standard accumulation equation $k_{jt+1} = (1-\delta)k_{jt} + x_{jt}$. Innovation expenditures $A_t z_{jt} i_{jt}$ yield a higher probability of a successful innovation $\eta(i_{jt})$, which permanently raises productivity by a factor Δ:

$$\log z_{jt+1} = \begin{Bmatrix} \log z_{jt} + \Delta + \varepsilon_{jt+1} \text{ with probability } \eta(i_{jt}) \\ \log z_{jt} + \varepsilon_{jt+1} \text{ with probability } 1-\eta(i_{jt}) \end{Bmatrix}, \tag{1}$$

where $\varepsilon_{jt+1} \sim N(0, \sigma_\varepsilon)$ are idiosyncratic shocks to productivity growth unrelated to innovation. Note that the innovation expenditures $A_t z_{jt} i_{jt}$ required to produced a particular probability of success are increasing in firm-level productivity z_{jt}. This assumption captures the notion that ideas become harder to find over time because the easy ideas are "fished out" (see e.g., Jones (1995) and the evidence in Bloom et al. (2020)).

We assume that firms cannot sell their existing ideas, i.e. $i_{jt} \geq 0$. This assumption reflects the numerous frictions in the market for ideas and is a parsimonious way to capture periods of inactivity in innovation (which are common in the data). Section 8 describes how we can relax this assumption by allowing for a frictional market for ideas.

Firms have two sources of finance for their investment and innovation expenditures, both of which are subject to frictions. First, firms can borrow externally, but this borrowing is

[11] The endogenous growth literature typically assumes constant returns to scale in objects (here capital) on the basis of the replication argument, which then implies increasing returns to scale in objects and ideas jointly (here capital and productivity). Our model is consistent with this view if we alternatively interpret decreasing returns as reflecting a downward-sloping demand curve, which allows the fundamental production function to be constant returns to objects.

[12] Exit shocks are a common tool in the financial frictions literature to ensure that firms do not permanently outgrow their financial frictions in the long run (see, e.g., Khan and Thomas (2013)).

subject to the collateral constraint $b_{jt+1} \leq \theta k_{jt+1}$.[13,14] Second, firms can use their internal resources, but they cannot raise new equity. This assumption implies that dividend payments must be nonnegative:[15]

$$d_{jt} = (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha} + (1-\delta) k_{jt} - b_{jt} - k_{jt+1} - A_t z_{jt} i_{jt} + \frac{b_{jt+1}}{1+r_t} \geq 0.$$

Finally, following Romer (1990), we capture the non-rivalry of ideas by assuming that aggregate productivity is

$$A_t = \left(\int z_{jt} dj \right)^{a}. \tag{2}$$

This assumption captures the idea that firms can observe some average level of accumulated ideas in the economy, but only a fraction $a \geq 0$ is relevant for their own production decisions.

Households There is a representative household with preferences represented by the utility function $\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma}-1}{1-\sigma}$, where $1/\sigma$ is the elasticity of intertemporal substitution (EIS). Since there is no aggregate uncertainty, firms discount future profits using the implicit risk-free rate

$$\frac{1}{1+r_t} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma}. \tag{3}$$

Summary of Key Features There are four key differences between capital k and ideas z in our model:

(i) *Technological risk*: capital accumulation is deterministic but idea arrival is risky, gov-

[13]This constraint can be derived from an environment in which firms lack commitment to repay their debts, and lenders can seize a fraction $\frac{\theta}{1-\delta}$ of their capital if firms default. If firms do not face other costs from default and lenders are only willing to offer risk-free debt contracts, this means firms will face the constraint $b_{jt+1} \leq \theta k_{jt+1}$ (for additional details, see Ottonello, Perez and Varraso, 2022). This formulation implicitly assumes that lenders cannot recover any of the ideas generated by the firm, z_{jt}. This assumption can be rationalized if, due to the non-rivalry of ideas, the manager can freely copy existing ideas in default.

[14]Recent empirical evidence argues that a substantial fraction of corporate debt is collateralized by the value of firms' earnings rather than the value of their capital stock (see, e.g., Lian and Ma (2021)). In Section 8, we argue that allowing for earnings-based constraints would amplify our main results.

[15]We assume no equity issuance in order to parsimoniously capture equity issuance costs in our general equilibrium environment. At face value, the no-equity issuance constraint may seem overly strong, especially in light of the role that venture capitalists play in financing startups in the technology sector. In Section 4, we show that it is only the combined effects of the no-equity issuance and collateral constraints that matter for firms' investment and innovation decisions. Our calibration strategy disciplines the combined strength of these frictions, and we show that the implied frictions provide a good fit to the response of both investment and innovation to investment tax shocks.

erned by the arrival of new ideas $\eta(i_{jt})$.

(ii) *Collateralizability*: capital is collateralizeable in the borrowing constraint $b_{jt+1} \leq \theta k_{jt+1}$, but ideas are not collateralizable.

(iii) *Sellability*: capital is sellable, but ideas are not.

(iv) *Rivalry*: capital is fully rival but ideas are partially non-rival due to imitation from new entrants and through aggregate productivity A_t if $a > 0$.

3.2 Equilibrium

In order to define the equilibrium, it is convenient to formulate firms' decisions recursively. The firm's individual state variables are its individual productivity z_{jt} and its net worth $n_{jt} = (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha} + (1-\delta)k_{jt} - b_{jt}$. Exiting firms set $k_{jt+1} = b_{jt+1} = 0$, while continuing firms' decisions are characterized by the Bellman equation

$$v_t^{\text{cont}}(z,n) = \max_{k',i,b'} n - k' - A_t z i + \frac{b'}{1+r_t} + \frac{1}{1+r_t}\mathbb{E}_t\left[v_{t+1}(z',n')\right] \text{ s.t. } d \geq 0 \text{ and } b' \leq \theta k', \quad (4)$$

where $\mathbb{E}_t\left[v_{t+1}(z',n')\right]$ integrates over the next period's exit shock, innovation success, and idiosyncratic productivity shocks. The implied decision rules induce a law of motion for the measure of firms, $\Phi_{t+1}(z,n) = T(\Phi_t; k'(\cdot), i(\cdot), b'(\cdot))(z,n)$.

A *competitive equilibrium* is a sequence of value functions $v_t(z,n)$; policies $k_t'(z,n)$, $i_t(z,n)$, and $b_t'(z,n)$; measure of firms $\Phi_t(z,n)$; real interest rate r_t; and aggregate productivity A_t such that (i) firms optimize and the associated policy functions solve (4); (ii) the evolution of $\Phi_t(z,n)$ is consistent with firm decisions; (iii) the real interest rate r_t is given by (3) with $C_t = \int (y_{jt} - (k_{jt+1} - (1-\delta)k_{jt}) - A_t z_{jt} i_{jt})\, dj$; and (iv) aggregate productivity is from (2).

Balanced Growth Path Because productivity grows over time, the limiting behavior of the model is characterized by a balanced growth path (BGP). The BGP is determined by the growth rate of $Z_t = A_t \times (\int z_{jt} dj) = (\int z_{jt} dz)^{1+a}$. In Appendix A, we show that this growth rate $1+g=\frac{Z_{t+1}}{Z_t}$ equals the growth rate of all macroeconomic aggregates and scales the individual decisions and measure of firms.

4 The Pecking Order of Firm Growth

We now study the economic mechanisms that govern firms' decisions to growth through investment and innovation. The marginal cost of pursuing either of these activities is given by the firm's *shadow value of funds* $\frac{\partial v_t(z,n)}{\partial n}$. This object represents the marginal value of keeping extra resources in the firm, and is therefore the opportunity cost of instead spending those resources on either investment or innovation. Appendix A shows that $\frac{\partial v_t(z,n)}{\partial n}=1+\lambda_t(z,n)$, where $\lambda_t(z,n)$ is the Lagrange multiplier on the nonnegativity constraint on dividends. Furthermore, $\lambda_t(z,n)$ equals the expected value of the multipliers on future collateral constraints in all possible states of the world. Hence, $\lambda_t(z,n)$ captures how financial frictions create a wedge between the household's value of funds, 1, and the firm's shadow value of funds, $1+\lambda_t(z,n)$. We therefore refer to $\lambda_t(z,n)$ as the firm's *financial wedge*.

Proposition 1 characterizes firms' decision rules, building on Khan and Thomas (2013)'s similar result in a model without innovation:

***Proposition* 1.** *Consider a firm in period t that will continue operations in $t+1$, has productivity z, and has net worth n. Then there exist two functions $\overline{n}_t(z)$ and $\underline{n}_t(z,n)$ that partition the individual state space:*

(i) ***Financially unconstrained****: If $n\geq\overline{n}_t(z)$, then the financial wedge $\lambda_t(z,n)=0$. In this case, the firm is indifferent over any combination of external financing b' and internal financing d leaves them financially unconstrained. As in Khan and Thomas (2013), we resolve this indeterminacy by requiring that firms pursue the "minimum savings policy," i.e., the smallest level of $b'\equiv b^*(z)$ that leaves them unconstrained with probability one. Being financially unconstrained is an absorbing state.*

(ii) ***Currently constrained****: If $n\leq\underline{n}_t(z,n)$, then both the collateral constraint binds $b'=\theta k'$ and the financial wedge is positive $\lambda_t(z,n)>0$.*

(iii) ***Potentially constrained****: Otherwise, the collateral constraint is not currently binding $b' < \theta k'$ but the financial wedge is positive $\lambda_t(z,n) > 0$.*

In any of these cases, the optimal choices for investment $k'_t(z,n)$, innovation $i_t(z,n)$, and external financing $b'_t(z,n)$ solve the system

$$1 + \lambda_t(z,n) = \frac{1}{1+r_t}\mathbb{E}_t\left[(MPK_{t+1}(z',k') + 1 - \delta) \times (1 + \lambda_{t+1}(z',n'))\right] + \theta\mu_t(z,n) \tag{5}$$

$$1 + \lambda_t(z,n) \geq \frac{\eta'(i)}{A_t z}\frac{1}{1+r_t}\left(\mathbb{E}_t[v_{t+1}(z',n'|\text{ success })] - \mathbb{E}_t[v_{t+1}(z',n'|\text{ failure })]\right), \quad = \text{ if } i > 0 \tag{6}$$

$$k' + A_t z i = n + \frac{b'}{1+r_t} \text{ if } \lambda_t(z,n) > 0; \text{ otherwise, } b'_t(z,n) = b^*_t(z), \tag{7}$$

where $MPK_{t+1}(z',k') = \alpha\left(\frac{A_{t+1}z'}{k'}\right)^{1-\alpha}$ is the marginal product of capital and $\mu_t(z,n)$ is the Lagrange multiplier on the collateral constraint.

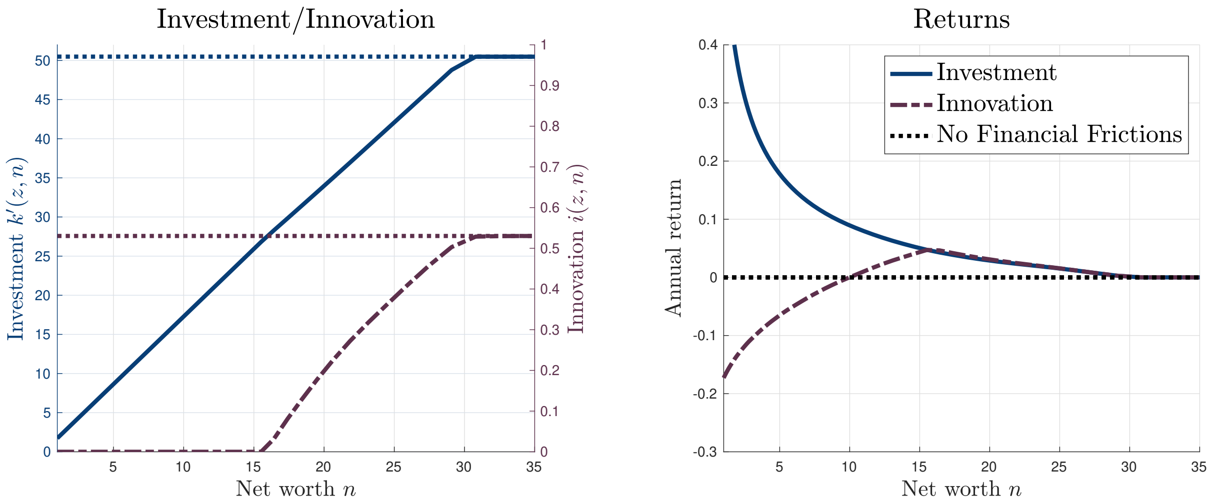
Proof. See Appendix A. ■

The first part of Proposition 1 describes three different regimes of financial constraints. *Financially unconstrained* firms have zero probability of facing a binding collateral constraint, which implies that their financial wedge is $\lambda_t(z,n) = 0$. These firms are able to follow the policy rules from the version of the model without financial frictions. The remaining two types of firms are affected by financial frictions in some way. *Currently constrained* firms' collateral constraint binds in the current period, which directly limits their ability to accumulate capital. *Potentially constrained* firms do not face a binding collateral constraint in the current period, but there is a positive probability of reaching a future state in which the constraint becomes binding. Financial frictions still affect these firms' decisions through precautionary motives.

The second part of Proposition 1 characterizes the decision rules of these three types of firms.[16] Equations (5) and (6) are the first-order conditions for investment and innovation. As discussed above, the marginal cost on the LHS of these equations is given by the shadow

[16] Our numerical algorithm solves the firm's problem by jointly iterating over the policy functions and the Lagrange multipliers $\lambda_t(z,n)$ in (the detrended version of) this system (5) - (7). This procedure is very fast because it avoids any numerical maximization or equation solving. In practice, we find computational runtimes comparable to using Carroll (2006)'s grid method, even though that method does not apply to this model. Our algorithm is applicable to other investment models in which the endogenous grid method does not apply, and therefore may be of interest to other researchers. See Appendix B for details.

FIGURE 2: The Pecking Order of Firm Growth



Notes: the left panel plots capital expenditures $k_t(z,n)$ (left axis) and innovation expenditures $i_t(z,n)$ (right axis) in market equilibrium BGP of the calibrated model for fixed z. The right panel plots the return to these activities, defined as the RHS of Euler equations (5) and (6) minus 1. "No financial frictions" refers to the model in which all firms following the unconstrained policies $k^*(z)$ and $i^*(z)$ from Proposition 1, but using the same real interest rate from the market BGP.

value of funds $1 + \lambda_t(z,n)$. The marginal benefit of capital on the RHS of (5) is given by two terms: the (discounted) expected marginal product of capital in the next period and the marginal collateral benefit provided by additional capital. This first-order condition always holds with equality because firms can freely sell capital. In contrast, because firms face a non-negativity constraint on innovation, the innovation first-order condition may not hold with equality if $i_t(z,n) = 0$ is the optimal solution. The marginal benefit of innovation on the RHS of (6) is the marginal improvement in the probability of success per unit of innovation expenditure times the expected improvement in firm value from a successful innovation. Equation (7) is the nonnegativity constraint on dividends, which binds as long as the firm has a positive financial wedge $\lambda_t(z,n) > 0$. In this case, innovation and investment expenditures must be financed out of either internal net worth or new borrowing.

Illustrating the Pecking Order We now use equations (5) and (6) to illustrate the key mechanisms governing the growth of a typical firm in the economy. Figure 2 plots the investment and innovation policies as a function of net worth for a firm with average productivity. These plots are generated using our calibrated model; see Section 5 for details of the calibration.

We say that these policies determine a *pecking order of firm growth* because there are three distinct regions in net worth space for a given level of productivity z. In the first region, for low values of net worth, the firm grows only through investment; the firm spends all of its available resources on capital and sets innovation expenditures to zero $i_t(z, n) = 0$. The capital accumulation of these firms is limited by their net worth, which implies that the return to capital is above the return to innovation. As net worth increases, the firm is able to accumulate more capital, which drives down its return due to the diminishing marginal product of capital. At the same time, higher capital also raises the return to innovation because TFP and capital are complements in production.

As net worth continues to increase, the firm enters the second region in the pecking order, in which it grows through both investment and innovation. In this region, the innovation Euler equation (6) holds with equality, so the returns to capital and innovation must be equalized. However, both (net) returns are still greater than zero because the financial wedge is positive $\lambda_t(z, n) > 0$. This fact implies that firms do not pay dividends (see equation (7)). In this case, investment and innovation are substitutes because, for a given level of net worth, higher investment must be accompanied by lower innovation and viceversa. On the other hand, a marginal increase in net worth will increase both investment and innovation, with the associated sensitivities determined by the slope of the expenditure curves.

Finally, for sufficiently high levels of net worth, the firm enters the last region of the pecking order in which it grows only through innovation. Conditional on its value of productivity z, a firm in this region has reached its optimal scale given its current level of productivity, $k^*(z)$. At this point, the financial wedge $\lambda_t(z, n) = 0$, driving the net returns to investment and innovation to zero and implying that firm's policies become independent of net worth. Hence, the only way in which the firm will grow further is the realization a successful innovation. If this happens, the firm's productivity z will jump up, at which point both returns will also jump up and the firm may re-enter the dynamics described above.

This pecking order of firm growth is consistent with empirical patterns of investment and innovation that we documented in Section 2. Smaller and younger firms (who are more likely to be financially constrained in our model) spend relatively more on investment and relatively less on innovation than older and larger firms (who are more likely to be financially unconstrained).

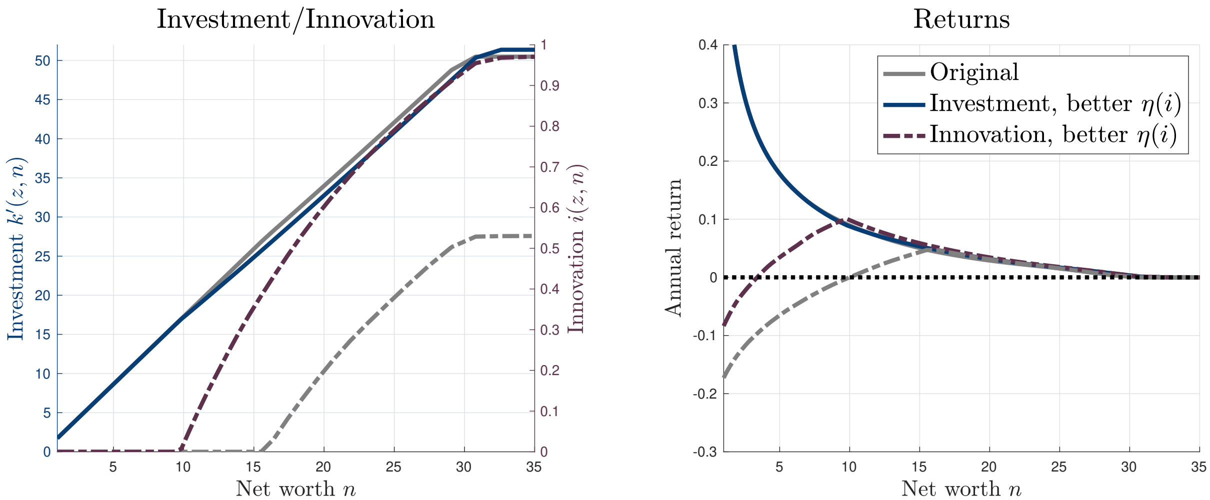
Key Forces Governing the Pecking Order While all parameters shape this pecking order of firm growth, we find that two key sets of parameters are particularly important: the degree of financial frictions θ and the efficiency of the innovation technology $\eta(i)$.

Without financial frictions, the model would not have a pecking order at all; firms would immediately be able to leverage up to their optimal scale given current productivity. As Figure 2 shows, the decision rules would be independent of net worth and the financial wedge would be $\lambda_t(z,n)=0$. In this case, investment and innovation become independent of size and age, inconsistent with the evidence presented in Section 2. Hence, financial frictions are the key model ingredient which allows us to be consistent with that motivating evidence.[17]

Figure 3 illustrates the effects of a more efficient innovation technology, which raises the success probability $\eta(i)$ for any level of innovation expenditures. The higher success probability shifts up the returns to innovation, which implies that it intersects the returns to capital at a lower level of net worth. Therefore, firms begin innovating and enter the second region of the pecking order for lower levels of net worth than in the baseline. The firm's innovation rate is also higher in this region, which forces it to (slightly) reduce its capital expenditures given the flow of funds constraint (7). However, once the firm becomes unconstrained, its level of capital accumulation is higher than in the baseline scenario because higher innovation increases the expected marginal product of capital.

[17] Of course, there may be other sources of sluggish adjustment, such as capital adjustment costs, that generate small/young firms have a higher return to capital than large/old firms. We prefer to study the role of financial frictions for two reasons. First, there is a large literature which shows that financial frictions are an important determinant of investment dynamics (see, for example, Cooley and Quadrini, 2001; Khan and Thomas, 2013; Ottonello and Winberry, 2020). Second, financial frictions imply that investment and innovation are substitutes for constrained firms, which adjustment costs would not generate. This substitutability generates the policy tradeoff for policy intervention discussed in Section 7.

Figure 3: Role of Innovation Technology in the Pecking Order



Notes: the left panel plots capital expenditures $k_t(z,n)$ (left axis) and innovation expenditures $i_t(z,n)$ (right axis) in market equilibrium BGP of the calibrated model for fixed z. The right panel plots the return to these activities, defined as the RHS of the Euler equations (5) and (6) minus 1. "Better $\eta(i)$" refers to the model with higher η_0 than in our baseline calibration (see Section 5), but using the same real interest rate from the market BGP.

5 Parameterization

We calibrate the model to ensure that this pecking order of firm growth is in line with key features of the data. Section 5.1 describes the moments we use to discipline the key parameters described above. Section 5.2 uses these moments (and others) to calibrate the model. Finally, Section 5.3 shows that the model matches various untargeted statistics, including the response of investment and innovation to investment tax shocks.

5.1 Strategy for Disciplining Key Forces

Following the literature, we will choose the tightness of the collateral constraint θ to match the average leverage of firms in the data. Therefore, the main challenge in our calibration is to pin down the properties of the innovation technology: (i) the probability of a successful innovation $\eta(i)$ and (ii) the size of successful innovations Δ.

While we can arguably measure innovation in the data using R&D expenditures, there is no direct measure of its output (successful innovations).[18] One option is to directly measure the model's implied output—firm-level productivity—but that is problematic because of the well-known difficulties in estimating production functions and the inability to separate price from quantity for an economy-wide sample of firms. Another option is to proxy for successful innovations using patents, but patents cover a (potentially small) subset of all innovations and are difficult to convert back to units of productivity.

Given these difficulties, our approach is to infer successful innovations from firms' investment behavior rather than attempt to measure successful innovations directly. This approach holds promise because the long duration of capital assets makes investment an especially forward-looking decision. In our model, firms that receive a successful innovation experience an *investment spike*—a large but short-lived surge in their investment rate. Therefore, the responsiveness of investment spikes to R&D expenditures is informative about the innovation technology.

We study the relationship between investment spikes and R&D expenditures in our Compustat sample. Following Cooper and Haltiwanger (2006), we define investment spikes as periods in which a firm's investment rate is above 20%. In our sample, the frequency of investment spikes is 20% and the average size of an investment spike is 31.5%, similar to Cooper and Haltiwanger (2006)'s Census sample. The average R&D-to-sales ratio is 2.9%, and its standard deviation 5.7%. Among firm-years with positive R&D spending (45% of observations), the average R&D-to-sales ratio is 6.4%, in line with the ratio reported by Acemoglu et al. (2018) for the "continuously-innovative firms" in their Census sample. Appendix Table 6 presents a set of descriptive statistics for these variables and Appendix Figure 8 histograms of their distributions.

Innovation and investment spikes We estimate the linear probability model

$$\mathbb{1}\{\frac{x_{jt}}{k_{jt}} \geq 0.2\} = \alpha_j + \alpha_{st} + \sum_{h=1}^{H} \beta_h \left(\frac{\widehat{i}_{jt-h}}{\tilde{y}_{jt-h}} \right) + \Gamma' X_{jt} + \epsilon_{jt}, \tag{8}$$

[18] R&D expenditures may understate true innovation activity, especially for small firms with less developed accounting infrastructures. However, this concern is smaller in the context of our Compustat sample, which only includes large, publicly traded companies. In any event, we are not aware of a better-measured alternative to innovation in the data.

Table 1

Investment spikes and Innovation

	(1)	(2)	(3)
$\frac{i_{jt-1}}{\tilde{y}_{jt-1}}$	1.29	1.10	1.12
	(0.16)	(0.15)	(0.15)
$\frac{\text{cf}_{jt}}{k_{jt}}$		0.12	0.11
		(0.02)	(0.02)
years since spike$_{t-1}$			0.003
			(0.0008)
$\frac{k_{jt}}{n_{jt-1}}$			-0.016
			(0.003)
Observations	53,577	53,577	53,577
Adj. R^2	0.261	0.281	0.282

Notes: Results from estimating alternative versions of $\mathbb{1}\{\frac{x_{jt}}{k_{jt}} \geq 0.2\} = \alpha_j + \alpha_{ts} + \sum_{h=1}^{4} \beta_h \left(\frac{\widehat{i}_{jt-h}}{\tilde{y}_{jt-h}}\right) + \Gamma' X_{jt} + \varepsilon_{jt}$, where $\frac{x_{jt}}{k_{jt}}$ denotes the investment rate of firm j in period t; $\frac{\widehat{i}_{jt}}{\tilde{y}_{jt}}$ the R&D-to-sales ratio; α_j and α_{ts} firm and time by sector fixed effects; X_{jt} is a vector of firm-level controls; and ε_{jt} is a random error term. Column (1) reports estimates for a specification without including-firm level time-varying controls; Column (2) those that include cash flows ($\frac{\text{cf}_{jt}}{k_{jt}}$) as a control; and Column (3) those that also include the lumpy-investment controls (years since the last investment spike, years since spike$_{t-1}$, and the standardized capital-output ratio, $\frac{k_{jt}}{n_{jt-1}}$). To estimate the models reported in Columns (1) and (2), we restrict the sample to that with available observations in Column (3). For variable definitions and descriptive statistics, see Appendix C.

where $\frac{x_{jt}}{k_{jt}}$ denotes the investment rate of firm j in period t; $\frac{\widehat{i}_{jt}}{\tilde{y}_{jt}}$ the R&D-to-sales ratio; α_j and α_{st} firm and time by sector fixed effects; X_{jt} is a vector of firm-level controls; and ϵ_{jt} is a random error term. Our coefficient of interest, β_1, measures how the probability of an investment spike is related to its previous R&D expenditure. The vector of controls X_{jt} includes the ratio of cash flows to lagged capital—to absorb the effect that changes in firms' cash-on-hand has on both investment spikes and R&D expenditure—and two variables that the lumpy investment literature identifies as predictors of investment spikes: the number of years since the previous spike and the capital-employment ratio. Appendix C.1 details the construction of these variables. We set $H = 4$ for our baseline model and explore alternative lags in robustness analysis. We cluster standard errors two ways to account for correlation within firms and within years.

Table 1 shows that R&D expenditures are a powerful predictor of investment spikes. Column (1) reports the estimated coefficient β_1 from (8) without any additional controls

X_{jt}. Quantitatively, the estimated coefficient implies that have an R&D-to-sales ratio one standard deviation above the mean increases the probability of an investment spike by 7pp, i.e. a nearly 40% increase in the probability of a spike. Column (2) shows that estimate survives controlling for changes in cash flow which may independently affect both investment and R&D expenditures. Column (3) shows that the coefficient is virtually unaffected by controlling for the years since the last spike and the capital to labor ratio, which are important summaries of the incentives to invest in fixed cost models.

While we do not have exogenous variation in R&D expenditures to identify the causal effect of innovation on investment spikes, we view these results are suggesting a tight link between the two. Therefore, we will target the estimated coefficient β_1 in our model calibration by running the same regression on model-simulated data. Appendix C.2.3 presents robustness analysis and additional supportive evidence on this tight relationship between innovation and investment spikes. Appendix Table 7 shows that the results presented in Table 8 are robust alternative model specifications: an alternative definition of investment spikes (that considers a sector-level threshold instead of an absolute threshold), alternative lags of R&D-to-sales ratios, and additional controls used in the investment literature (e.g., size, sales growth, and the share of current assets). Complementing this evidence, Appendix Table 13 presents an event study, showing that R&D-to-sales tend to increase around investment-spike episodes.

5.2 Calibration

With this evidence in hand, we now calibrate the model. We proceed in two steps. First, we fix a subset of parameters to match standard aggregate targets. Second, we choose the remaining parameters to match moments in the data.

Table 2 contains the parameters that we exogenously fix. We set the EIS $1/\sigma = 1$, implying log utility. We set the household's discount factor $\beta = 0.98$ so that the real interest rate is 4% annually along the BGP. We set the elasticity of output with respect to inputs to be $\alpha = 0.55$, the value Hennessy and Whited (2005) estimate for Compustat firms and close to the 0.59 value Cooper and Haltiwanger (2006) estimate for manufacturing plans.[19] We set

[19]In the model with labor, α would be the total elasticity of output with respect to all inputs, consistent

TABLE 2
FIXED PARAMETERS

Parameter	Description	Value
Household		
β	Discount factor	0.98
$1/\sigma$	EIS	1.00
Firms		
α	Output elasticity w.r.t inputs	0.55
δ	Depreciation rate	0.08
π_d	Exit rate	0.08

Notes: parameters chosen exogenously to match external targets.

the depreciation rate to $\delta = 8\%$ annually to imply an aggregate investment-to-capital ratio of 10% along the BGP. Finally, we assume $\pi_d = 8\%$ of firms exit per year, broadly consistent with exit rates in both the Business Dynamics Statistics (BDS) and in our Compustat sample.

Table 3 contains the endogenously chosen parameters and the moments we target in the data. For targets drawn from Compustat data, we mirror the sample selection into Compustat by conditioning on firm age. The first three parameters govern the innovation technology: η_0 controls the efficiency of the success probability $\eta(i)$, Δ is the size of successful innovations, and a controls the non-rivalry of ideas in A_t. The corresponding first three moments in the right panel of the table contain the intuitively strongest source of identifying variation in the data (though of course all parameters are jointly chosen to match all targets). As discussed above, the regression coefficient of the probability of investment spikes on lagged R&D spending from Table 8 pins down the efficiency of the success probability $\eta(i)$. The average size of the resulting investment spikes then pins down the size of successful innovations Δ. Given this innovation technology, we then infer the degree of non-rivalry of ideas among incumbent firms, a, to match a long-run growth rate of 2% per year.

The most natural point of comparison for our estimated innovation technology is to the empirical literature on the response of patenting to R&D spending, which is often used to discipline creative destruction models (see, e.g., Acemoglu et al. (2018)). These studies typically find an average elasticity of successful innovation to R&D around 0.5, while our estimates imply an average elasticity 0.77. One interpretation of this finding is that invest-

with our calibration targets. See Footnote 10 for more details about the model with labor.

TABLE 3
FITTED PARAMETERS

Parameter	Description	Value	Target (all joint)	Data	Model
Innovation technology					
η_0	Efficiency	0.20	Regression coefficient	1.10	1.08
Δ	Size of innovations	0.20	$\mathbb{E}[x_{jt}/k_{jt}\vert\text{spike}]$	0.32	0.32
a	Non-rivalry of ideas	0.22	Growth rate	0.02	0.02
Financial frictions					
θ	Collateral constraint	0.45	$\mathbb{E}[b_{jt}/k_{jt}]$	0.13	0.15
Productivity shocks					
σ_ε	SD of shocks	0.02	$\sigma(x_{jt}/k_{jt})$	0.13	0.10

Notes: left panel contains the parameters chosen to match the moments in the right table. "Regression coefficient" is the regression coefficient on lagged R&D-to-sales in Table 1 column (2). $\mathbb{E}[x_{jt}/k_{jt}|\text{spike}]$ is the average size of investment spikes in our Compustat sample described in Section 5.1. "Growth rate" is the aggregate growth rate along the market equilibrium BGP. $\mathbb{E}[b_{jt}/k_{jt}]$ is the average (net) leverage of firms in our sample. Finally, $\sigma(x_{jt}/k_{jt})$ is the standard deviation of investment rates in our sample.

ment spikes capture a broader set of innovations than do patents, though it is not the only interpretation.

The remaining parameters are pinned down by standard moments. We choose the degree of financial frictions θ in order to match the average (net) leverage of firms in our Compustat sample. While this target may be problematic given that we do not model the tax advantage of debt, we show below that it implies realistic heterogeneity in the response of innovation to the Bonus Depreciation Allowance. We choose the dispersion of idiosyncratic shocks σ_ε to match the dispersion of investment rates in the data.

Finally, as discussed in Section 3, we choose the distribution of new entrants to capture the idea that new entrants imitate incumbent firms. We assume that productivity and capital are drawn independent log-normal distributions. We assume that the mean of the distribution of productivity equals the mean of the distribution of incumbent firms, and set the dispersion in those draws to $\sigma_z = \Delta$. We assume that the mean of the capital distribution is proportional to the mean of the incumbent distribution and set the constant of proportionality to imply that new entrants' capital is roughly 20% of incumbents' capital. We choose the dispersion of new entrants' capital to match the relative dispersion of new entrants to incumbents (also roughly 20%).

5.3 Validation

We now show that the calibrated model matches untargeted statistics in the data. Importantly, the model matches new evidence on the response of innovation to exogenous changes in the after-tax price of investment, validating the link between investment and innovation that is at the core of our model.

Sources of Firm Heterogeneity Appendix D analyzes the two sources of firm heterogeneity in our calibrated model: lifecycle dynamics and productivity differences (due to either successful innovations or productivity shocks). Following the pecking order of firm growth from Section 4, most young firms start investment-intensive but become more innovation-intensive as they age. Increases in productivity raise the marginal product of capital and shadow value of funds $1 + \lambda_t(z, n)$, which induces firms to invest and borrow more but innovate less. These dynamics imply positive investment- and innovation-cash flow sensitivities, as in the data. We also show that the model matches a number of untargeted moments from our Compustat sample.

Investment Tax Shocks To study the response of innovation to changes in the incentives to invest, we exploit variation in the after-tax price of investment induced by the Bonus Depreciation Allowance. The Bonus is a countercyclical investment stimulus policy, used in the 2001 and 2008 recessions, which allows firms to deduct a fraction $b_t \in [0, 1]$ of investment expenses from their tax bill immediately (and apply the standard depreciation schedule to the remaining $1 - b_t$ fraction of expenditures). By bringing forward future tax deductions into the present, the policy increases the present value of tax deductions by $\Delta\widehat{\zeta}_t = b_t(1 - \widehat{\zeta})$ where $\widehat{\zeta} < 1$ is the present value of deductions under the baseline schedule.

Zwick and Mahon (2017) show that sectoral heterogeneity in the baseline tax depreciation schedule across sectors s provides exogenous variation that can be used to identify the effect of the Bonus on investment. Table 4 Column (1) replicates Zwick and Mahon (2017)'s estimates of the effect of the Bonus on investment in our Compustat sample. Specifically,

Table 4
Bonus Depreciation Allowance in the Data and the Model

	(1) $\frac{x_{jt}}{k_{jt}}$, data	(2) $\frac{x_{jt}}{k_{jt}}$, model	(3) $\frac{i_{jt}}{y_{jt}}$, data	(4) $\frac{i_{jt}}{y_{jt}}$, model	(5) $\frac{i_{jt}}{y_{jt}}$, small	(6) $\frac{i_{jt}}{y_{jt}}$, large
$\frac{1-\tau PV_{st}}{1-\tau}$	-1.41	-1.51	-0.24	-0.31	-0.53	-0.13
	(0.25)		(0.06)		(0.23)	(0.05)
R^2	0.41	0.78	0.89	0.52	0.85	0.91

Notes: estimates of $\widehat{\gamma}$ from the regression (9) in columns (1) and (2) or from the regression (10) in columns (3) - (6). Standard errors are two-way clustered by firms and years. "Model" columns (2) and (4) replicate the regressions on model-simulated data in response to a shock equivalent to a 50% bonus depreciation allowance which reverts back to its long-run average following an AR(1) process with annual persistence 0.75 (giving a half-life of roughly two years). "Small" firms in column (5) are those whose average sales are in the bottom 3 deciles of the sales distribution. "Large" firms in column (6) have average sales in the top 3 deciles of the sales distribution.

we estimate the regression

$$\frac{x_{jt}}{k_{jt}} = \alpha_i + \alpha_t + \gamma \frac{1 - \tau\widehat{\zeta}_{st}}{1-\tau} + \Gamma' X_{jt} + \epsilon_{jt}, \tag{9}$$

where α_i is a firm fixed effect, α_t is a time fixed effect, X_{jt} is a vector of controls, and ϵ_{jt} are residuals. In the data, we estimate the regression coefficient $\widehat{\gamma} = -1.41$, which is close to Zwick and Mahon (2017)'s estimate of -1.53 using high-quality IRS microdata. A 50% bonus would increase increased the average value of $\frac{1-\tau PV_{st}}{1-\tau}$ by -0.03, implying its direct effect increased the average firm's investment rate by $-0.03 \times -1.41 = 0.04$ compared to its unconditional average of 0.14.

Column (2) in Table 4 shows that the model's implied regression coefficient is $\widehat{\gamma} = -1.51$, very close to the data. We replicate the Bonus shock in our model by feeding in a decline in the relative price of investment equivalent to a 50% bonus.[20] We interpret this finding as validating the degree of investment frictions in the model.

Column (3) in Table 4 documents a new empirical finding: the Bonus also substantially raises innovation expenditures. We estimate the regression

$$\frac{\widehat{i}_{jt}}{\widetilde{y}_{jt}} = \alpha_i + \alpha_t + \gamma \frac{1 - \tau\widehat{\zeta}_{st}}{1-\tau} + \Gamma' X_{jt} + \epsilon_{jt}, \tag{10}$$

[20] Appendix F shows that the Bonus is isomorphic to a temporary shock to the relative price of capital in our model. We assume that the shock mean-reverts with a half-life of two years. Since the empirical specification (10) includes time fixed effects which absorb general equilibrium effects, we keep the real interest rate fixed at its initial value $r_t = r^*$ for this exercise.

which replaces the investment rate on the RHS of (9) with the RD-to-sales ratio $\widehat{i}_{jt}/\widetilde{y}_{jt}$. Note that the denominator $\widetilde{y}_{jt}$ is lagged sales in the past five years, so it is predetermined in the period of the shock. Quantitatively, this estimated coefficient implies that a 50% bonus directly raises the average firm's RD-to-sales ratio by about 0.8pp relative to its unconditional average of 2.9pp — a nearly 30% increase in innovation expenditures. Column (4) shows that our model matches the empirical response of innovation to the bonus nearly exactly, validating the link between investment and innovation that is at the core of our framework.

Appendix D Figure 18 shows that the model's cross-price elasticity of innovation with respect to investment $\frac{\partial \log i(z,n)}{\partial \log(1-\tau\zeta_t)}$ is positive but heterogeneous across firms. Unconstrained firms have a positive elasticity because higher investment also raises the return to innovation due to the complementarity between capital and productivity. On the other hand, constrained firms have a positive elasticity because the shock lowers their after-tax expenditures on investment, freeing up cash flows to finance innovation. Quantitatively, this cash flow channel is larger than the complementarity channel, consistent with the central role of financial frictions in our analysis.

Columns (5) and (6) in Table 4 confirm these heterogeneous responses in the data, providing further validation of the role of financial frictions in linking innovation and investment. Following Zwick and Mahon (2017), we define small firms as those whose average sales are in the bottom three deciles of the distribution and large firms whose sales are in the top three deciles. Small firms' innovation expenditures are about five times as responsive to the bonus as are large firms, consistent with the heterogeneous responses in our model.

6 The Impact of Financial Frictions on Economic Growth

The pecking order of firm growth from Section 4 already showed that financial frictions delay innovation at the firm level. Our goal in this section is to understand the implications of this finding for the aggregate economy. Along the BGP, the aggregate growth rate is

$$g \approx (1+a)(e^{\Delta}-1)\int \eta(i_{jt})dj. \tag{11}$$

Table 5
Role of Financial Frictions in Long-Run Growth

	Growth rate g	$\mathbb{E}[r^k] - r$	$\mathbb{E}[r^i - r \mid i > 0] - r$	$\sigma(r^k - r)$	$\sigma(r^i - r)$
Calibrated model	2.0%	2.3%	0.5%	4.8%	3.0%
No financial frictions (fixed r)	3.4%	0	0	0	0
No financial frictions (GE)	2.13%	0	0	0	0

Notes: cross-sectional statistics from the stationary distribution in the market equilibrium BGP. "Calibrated model" refers to full model calibrated as in main text. "No financial frictions (fixed r)" refers to model in which all firms follow the unconstrained policies $k^*(z)$ and $i^*(z)$ from Proposition 1, and the real interest rate is the same as in the calibrated model. "No financial frictions (GE)" refers to the same model in which the real interest rate is consistent with the new level of consumption growth. $r^k - r$ is the return to capital from the RHS of (5), expressed as an excess return over the risk-free rate. r^i is the return to innovation from the RHS of (6), also expressed as an excess return over the risk-free rate.

Hence, to the extent that financial frictions lower the average probability of receiving a new idea, they will lower economic growth.

Aggregate Effects We compute the effects of financial frictions by comparing our calibrated model to a version of the model in which there are no financial frictions (so that all firms follow the unconstrained policies $k^*(z)$ and $i^*(z)$ from Proposition 1). We perform this comparative static in two steps. First, we hold the real interest rate fixed at its initial level in order to calculate the direct effect of removing financial frictions. Table 5 shows that the aggregate growth rate would increase to 3.4% per year in this case — nearly double the growth rate in the calibrated model. Second, we allow the real interest rate to adjust up in line with the higher consumption growth rate. The higher real interest rate reduces the returns to innovation, dampening the increase in the overall growth rate. In this case, the annual growth rate increases to 2.13% per year, i.e. 13bps per year higher than in the calibrated model. The strength of this force depends on the EIS; with an EIS of 2, for example, the growth rate without financial frictions would be nearly 2.25% per year.[21]

This exercise shows that an important cost of financial frictions on the economy is lower economic growth. This finding complements the existing literature which studies how financial frictions may distort the allocation of capital and, therefore, lower the level of TFP (e.g.

[21] We find the comparative static with respect to the EIS informative because interest rates are very powerful in influencing the decisions of unconstrained firms (given that they face no adjustment costs). We conjecture that, in a richer model with adjustment costs, higher interest rates would not have such a dramatic effect on growth, yielding results closer to the high EIS parameterization.

Buera, Kaboski and Shin (2011) or Midrigan and Xu (2014)). While that misallocation also occurs in our model, over long horizons the growth effects account for the majority of the costs of financial frictions.

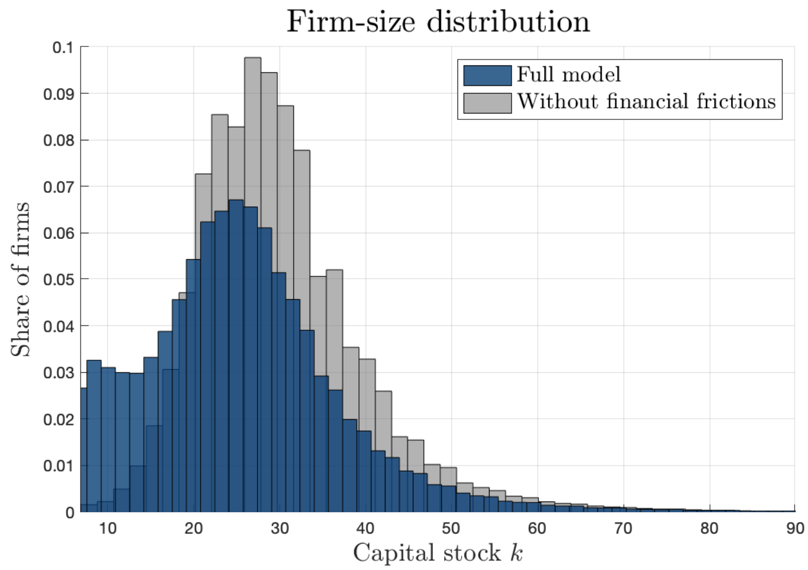
Table 5 also shows how financial frictions distort both the total amount of investment and innovation as well as its allocation across firms. Specifically, we study the excess return to capital or innovation from the RHS of the Euler equations (5) and (6), relative to the risk-free rate. Without financial frictions, all of these excess returns would be identically zero across all firms because they have no financial wedge $\lambda_t(z, n) = 0$. Financial frictions imply average excess return to capital and innovation are 2.3% and 0.5%, respectively. In addition, the "misallocation" implied by the dispersion of returns is higher for capital (dispersion of 4.8%) compared to innovation (with dispersion of 3.0%).

Distribution of Innovation Financial frictions also affect the distribution of innovation across firms; in our calibrated model, the majority of innovation (86%) is performed by large, unconstrained firms. This force thickens the right tail of the firm size distribution relative its mode. We illustrate this mechanism in Figure 4, which compares the distribution of detrended capital stocks in the two BGPs; given the differences in growth rates, direct comparisons across the two economies are not valid, but comparisons within each economy are still meaningful. From this perspective, the size distribution in our full model has more mass in both the left and right tails than does the distribution without financial frictions. The thickness of the left tail reflects the fact that it takes new entrants longer to grow, while the thickness of the right tail reflects the fact that unconstrained firms who survive follow a random growth process with exogenous death.[22]

Appendix G shows that temporary financial shocks θ_t do not have a particularly persistent effect in our model (despite the sizeable effects of *permanent* differences in θ described above). This result contrasts with the stylized fact that financial shocks have more persistent negative

[22]In fact, these random growth with death dynamics in the right tail generate a Pareto tail (see, e.g., Jones and Kim (2018)). Unfortunately, the model's tail is thinner than in the data because the expected size of successful innovations must be relatively small to match the average size of investment spikes in the data. However, we can thicken the tail by incorporating heterogeneity in the size of successful innovations, which would create heterogeneity in the expected growth rates (again in the spirit of Jones and Kim (2018)). In this extension, the average of these growth rates would still be pinned down by the average size of investment spikes, but the thickness of the right tail would be driven by firms with higher realized growth.

FIGURE 4: Firm-Size Distribution



Notes: distribution of capital along the balanced growth path. Capital stocks have been detrended in order to compute a stationary distribution, but the resulting distribution has the same cross-sectional properties as the raw distribution (see Appendix A). "Full model" refers to our calibrated model. "Without financial frictions" refers to the version of the model in which firms follow the unconstrained policies $k^*(z)$ and $i^*(z)$ from Proposition 1.

economic effects in the data (e.g., Cerra and Saxena, 2008). Based on representative firm models, some have argued that this persistence is driven by tighter financial constraints reducing innovation and therefore medium-term growth. However, in our heterogeneous firm model, the majority of innovation at a given time is performed by unconstrained firms, as described above. These firms are not directly affected by the shock and therefore face no impulse to lower innovation; in fact, since the real interest rate falls in general equilibrium, aggregate innovation *rises* following the shock.

7 Policy Implications

The competitive equilibrium studied above will generally not be socially efficient because firms do not internalize the non-rivalry of the ideas they produce, motivating welfare-improving policies. In order to better understand the implications of this externality, Section

7.1 studies the problem of a constrained-efficient planner's problem who internalizes the externality. Section 7.2 uses these results to study how two commonly-used policies, innovation subsidies and investment tax cuts, address the non-rivalry of ideas.

7.1 Financial Frictions and the Non-Rivalry of Ideas

In principle, our model actually has two externalities which may motivate policy intervention: the non-rivalry of ideas, which implies the private returns to innovation lie below the social returns, and pecuniary externalities through the real interest rate, which may distort firms' decisions due to financial frictions. We focus on the positive non-rivalry externality, which is new to our paper. In order to more precisely understand how this externality affects the allocation of investment and innovation, we characterize the problem of a constrained-efficient social planner who faces the same financial constraints as private firms but internalizes the non-rivalry of ideas. In order to abstract from the pecuniary externalities, we assume that the planner takes the borrowing rate as given.[23] However, the planner does take into account its effects on the marginal rate of substitution across time, and we ensure that the borrowing rate is equal to this marginal rate of substitution in an outer loop of the procedure.[24]

Planner's Problem Appendix E formulates the planner's problem recursively. The problem is technically challenging because the state variable is the entire distribution of firms and the control variables are entire functions of the firms' individual states. We overcome this challenge by solving the planner's problem in the function space following Lucas and Moll (2014) and Nuño and Moll (2018), who solve the planner's problem using Gâteaux derivatives (which are the natural generalization of partial derivatives in the function space). See Appendix E for details. We arrive at the following natural characterization of the planner's

[23]These pecuniary externalities have been extensively studied in the literature (see, for example Geanakoplos and Polemarchakis, 1986; Lorenzoni, 2008; Dávila and Korinek, 2018). That said, our model does potentially generate an interaction between innovation decisions and misallocation through this pecuniary externality. See Section 8 for more details.

[24]We also assume that the planner takes as given the distribution of new entrants, i.e. does not take into the positive externality through imitation. We make this simplifying assumption because we take the entry process into our model as exogenous; incorporating this margin would only further increase the positive externality of innovation. While we have extended our characterization of the constrained-efficient allocation taking these externalities into account, we are still working on numerically implementing those changes. Results available upon request.

allocation:

Proposition 2. *In the constrained-efficient equilibrium, individual allocations solve the augmented Bellman equation*

$$\omega_t^{cont}(z,n) = \max_{k',b',i} n - k' - A_t z i + \frac{b'}{1+r_t} + \Lambda_t z + \frac{1}{1+r_t}\mathbb{E}_t[\omega_{t+1}(z',n')] \ s.t.\ d \geq 0 \ and \ b' \leq \theta k' \tag{12}$$

where Λ_t is the planner's shadow value of the non-rivalry externality:

$$\Lambda_t = \left[a\left(\int z_{jt}dj\right)^{a-1}\right] \times \left[\int (1+\lambda_{jt})\left((1-\alpha)A_t^{-\alpha}z_{jt}^{1-\alpha}k_{jt}^{\alpha} - z_{jt}i_{jt}\right)dj\right]. \tag{13}$$

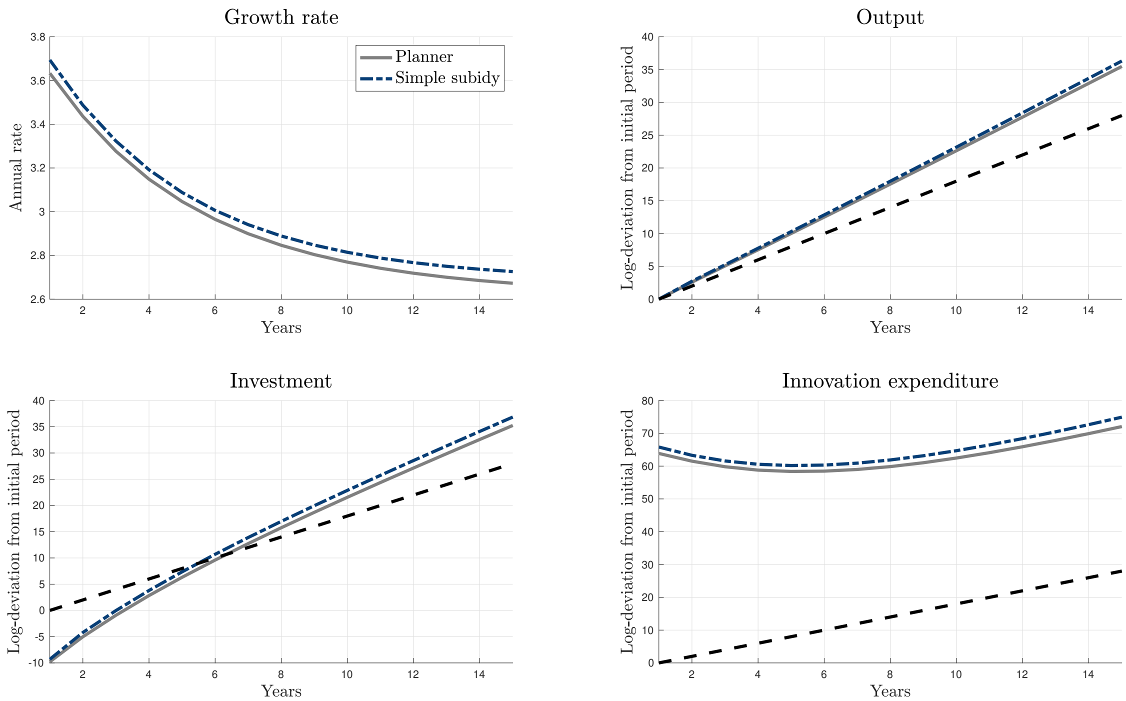
Proof. See Appendix E. ■

The only difference between the private Bellman equation (4) and the planner's augmented Bellman equation (12) is the shadow value of the non-rivalry externality, Λ_t. Equation (13) shows that this shadow value is the product of two terms: the marginal impact of an individual firm's productivity, z_{jt}, on aggregate productivity times the marginal social benefit of higher aggregate productivity. If this marginal benefit is positive, which we will find numerically, then private firms under-value innovation relative to the planner. This object is itself an integral of a product of two firm-level objects: the marginal increase in production net of innovation costs, $(1-\alpha)A_t^{-\alpha}z_{jt}^{1-\alpha}k_{jt}^{\alpha} - z_{jt}i_{jt}$, times the firms' shadow value of funds, $1+\lambda_{jt}$.

Proposition 4 shows that financial frictions amplify the positive externality from the non-rivalry of ideas; all else equal, a higher shadow value of funds $1+\lambda_{jt}$ increases the social value of higher aggregate productivity through the product in (13). This amplification occurs because the higher production net of innovation costs raises firms' cash flows and therefore alleviates the no-equity issuance constraint. The planner values loosening this constraint using the associated Lagrange multiplier λ_{jt}.

Planner's Allocation The characterization of the planner's allocation in Proposition 4 contains two important implications that successful policies must address. First, the planner prefers higher innovation because private firms do not internalize the positive externalities

Figure 5: Aggregate Transition Paths in Constrained-Efficient Allocation



Notes: aggregate transition paths chosen by planner (grey lines) and generated by the simple 13% innovation subsidy (dashed blue lines). Growth rate in top rate is in percentage points per year. Aggregate output, investment, and innovation expenditures in the remaining panels are in log-deviations from initial period. Dashed black lines are the growth trajectory in the initial market BGP.

of their innovation ($\Lambda_t > 0$). In contrast, the planner faces a meaningful tradeoff in terms of investment. On the one hand, higher innovation requires less investment for constrained firms due to their flow-of-funds constraint, i.e. investment and innovation are substitutes for constrained firms. On the other hand, higher innovation incentivizes more investment for unconstrained firms due to the complementarity between TFP and capital in production, i.e. investment and innovation are complements for unconstrained firms. In order to characterize this tradeoff, we solve for the planner's allocation starting from the equilibrium BGP, which involves a transition path to the new planner's BGP.

Figure 5 shows that the planner's balance between the substitutability vs. complementarity of innovation and investment changes over the course of the transition. Early on, the substitutability dominates in the sense that aggregate investment falls. This result occurs for two reasons. First, more firms are financially constrained early in the transition, implying more firms are in the substitutable region of the state space illustrated above. Second, the planner requires especially high innovation early on in the transition, implying constrained firms need to substantially reduce their investment. The planner values high innovation early on because more firms are constrained, which amplifies the planner's shadow value of the non-rivalry externality Λ_t as described above.

Over time, the complementarity between investment and innovation begins to dominate in the sense that aggregate investment eventually increases above the level along the initial BGP. This occurs because higher innovation raises net worth, implying that more firms are unconstrained and therefore in the complementary region of the state space. In addition, the planner's desired innovation falls over time as the shadow value of the externality falls as well.

7.2 Evaluating Innovation Subsidies and Investment Tax Cuts

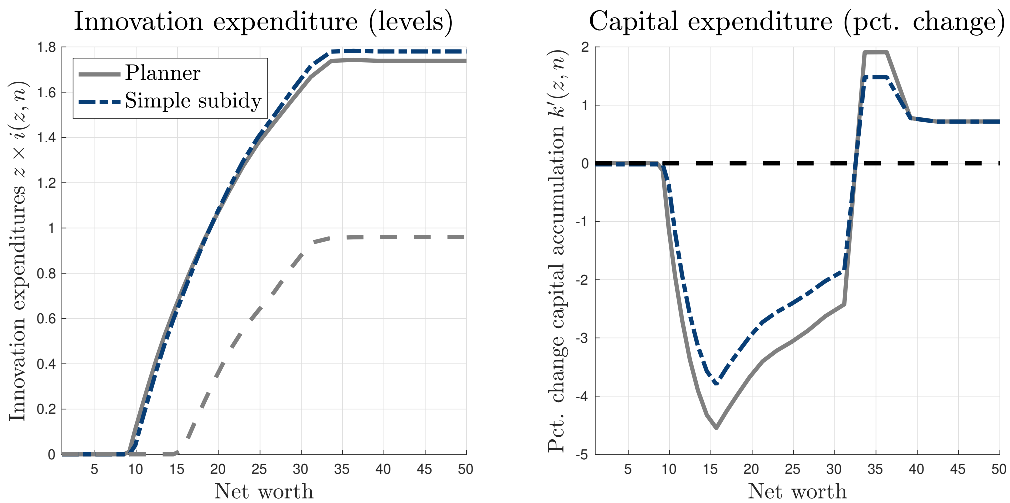
The planner's allocation seems difficult to implement in practice; in principle, one has to get both the allocation of innovation *and investment* correct, and the relevant tradeoffs vary across both firms and time.[25] In this section, we study how two simple, commonly-used policies perform with respect to these goals: an innovation subsidy and an investment tax cut.

Innovation Subsidy Given the subtle tradeoffs the planner balances, Figure 5 shows a striking finding: a simple innovation subsidy, constant across both firms and time, almost exactly replicates the aggregate paths chosen by the planner. We choose a 13% subsidy to generate the same long-run growth rate in the new BGP with the subsidy as the in the planner's BGP. Despite the fact that the subsidy is constant over time, the economy endogenously "front-loads" innovation early on in the transition, as desired by the planner.

Figure 6 shows that the innovation subsidy even gets the allocation of investment and innovation across firms approximately right. The left panel compares firms' policy rules for

[25] The planner's augmented Bellman equation (12) suggests one possible implementation: a time-varying transfer to firms proportional to individual productivity. This transfer would have to vary over time to mirror changes in the planner's shadow value Λ_t and vary across firms according to their productivity z_{jt}. Both of these objects are unobservable to policymakers in practice.

FIGURE 6: Planner's Decision Rules vs. Simple Subsidy, Initial Period



Notes: decision rules in the market equilibrium compared with the constrained-efficient allocation (grey lines) or to the policies under the simple innovation subsidy (blue lines). Left panel plots innovation expenditures $A_t z i_t(z,n)$ for a firm with average productivity. Right panel plots the percentage difference in capital accumulation policies.

innovation as a function of net worth in the market equilibrium, in the planner's allocation, and under the innovation subsidy. The simple subsidy also replicates the planner's desired changes to the innovation policy at the firm level; constrained firms endogenously increase their innovation by less than unconstrained firms according to the pattern of financial wedges $\lambda_t(z,n)$ across firms. The right panel shows that these firms also cut their capital expenditures in order to finance these higher innovation expenditures. However, the implied change in capital accumulation policies is not as close to the planner's policies as they are for innovation. Hence, it turns out that an appropriately-chosen innovation subsidy performs quite well, but not perfectly, in getting firms to internalize the non-rivalry externality.

Investment Tax Cut We now show that, to the extent that this nearly-optimal innovation subsidy is not fully available in practice, it can be partially substituted by an investment tax cut. Specifically, we find that cutting taxes on investment successfully increases innovation in the long run, but must also suboptimally increase investment in the short run.

We illustrate the connection between investment tax cuts and innovation using the Tax Cuts and Jobs Act (TCJA 2017) as an example. Appendix F extends our model to include a corporate tax code similar to the U.S. system. We show that the presence of this tax system changes the after-tax relative price of investment to be $1 - \tau\zeta_t$, where τ is the corporate tax rate and ζ_t is the present value of tax deductions per unit of investment. Before TCJA 2017, firms had to deduct investment expenditures over time according to the MACRS depreciation schedule, implying a present value $\zeta_t = \zeta^* < 1$ smaller than one due to discounting. The TCJA 2017 now allows firms to fully expense investment from their tax bill immediately, implying that $\zeta_t = 1$. We mirror this policy change in our model by studying a permanent decline in the after-tax price of investment by $-\tau(1 - \zeta^*)$.

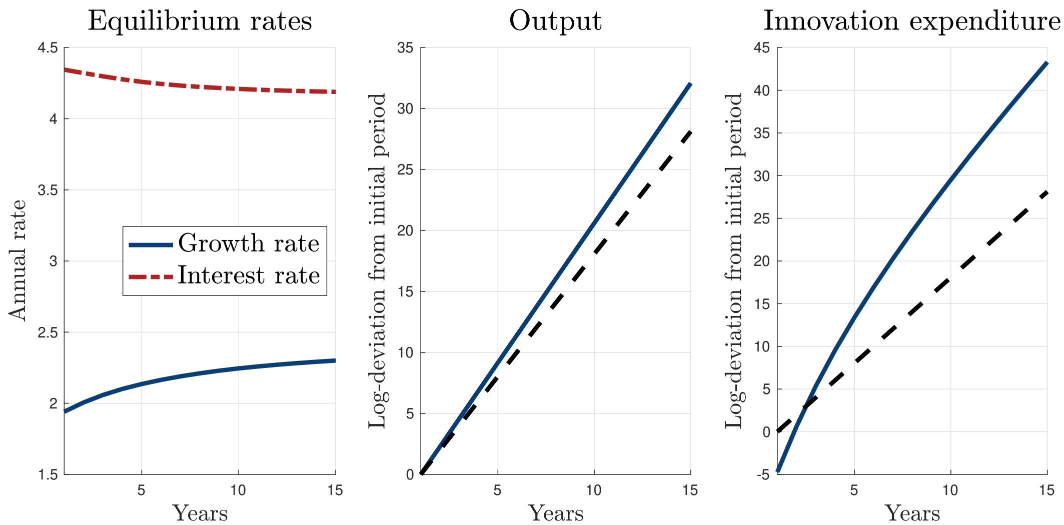
Figure 7 shows that, in our model, full expensing increases the long-run growth rate by nearly 20bps per year — a 10% increase in the annual growth rate of the economy. This result occurs for two reasons. First, for unconstrained firms, the complementarity of capital and TFP in production implies that the return to innovation increases with investment. Second, if after-tax capital expenditures fall, constrained firms can afford more innovation our of their current cash flows.

In contrast to our model, investment tax cuts would have no effect on the long-run growth rate in the neoclassical growth model. In the neoclassical model, cutting taxes on investment would increase the capital stock but, due to the diminishing marginal product of capital, that would only lead to an increase in the level of output, not its growth rate.

Figure 7 also shows that — despite raising innovation in the long run — full expensing *lowers* innovation in the first three years after its introduction. Innovation initially falls because the real interest rate r_t rises, lowering the expected returns to innovation. The real interest initially rises because full expensing increases capital demand, but capital supply is partially inelastic in the short run (due to consumption smoothing).[26] Over time, as capital supply catches up to the new long run level of demand, the real interest rate falls and the innovation rate rises. Hence, general equilibrium price effects are important for correctly capturing the dynamic effects of full expensing on innovation over time.

[26] Appendix G confirms this intuition by showing that innovation rises in the version of this experiment in which we keep the real interest rate fixed.

Figure 7: The Effects of Full Investment Expensing (TCJA 2017)



Notes: transition path following an unexpected, permanent decline in the relative price of capital $1-\zeta_t$ of the size equivalent to full expensing of investment, starting from the initial market BGP. Dashed lines correspond to the paths of investment, output, and innovation along the initial growth trajectory. Solid lines correspond to their actual paths in response to the change in the relative price of capital. Output and innovation expenditures expressed as log-deviations from initial period.

8 Extensions

We discuss four natural extensions of framework and how they would affect our main results.

Labor and a Negative Externality of Innovation The first extension we consider is to add labor as a variable input in production: $y_{jt} = (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha} \ell_{jt}^{\nu}$ with $\alpha + \nu < 1$. From the positive perspective, this extension would not significantly alter our main results; we would simply re-interpret our current production function as a reduced-form variable profit function after maximizing out labor, i.e. $\max_{\ell_{jt}} (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha} \ell_{jt}^{\nu} - w_t \ell_{jt}$. However, from the normative perspective, labor implies a negative pecuniary externality through the wage.

Appendix H shows that, in this extended model with labor, the planner now faces a tradeoff for socially optimal innovation. On the one hand, the non-rivalry of ideas provides a social incentive for higher innovation, as in our baseline analysis. But on the other hand, higher innovation will now also raise labor demand, which in turn raises labor costs and tightens financial frictions on constrained firms. Hence, optimal policy would have to balance the tradeoff between growth (coming from the non-rivalry of ideas) and misallocation (coming from these pecuniary externalities from the wage). Appendix H theoretically characterizes how the planner would balance this tradeoff, analogously to Proposition 4. However, numerically characterizing the solution requires adding an additional state variable to the firm's problem, so we leave it out of our baseline analysis for the sake of tractability.

Earnings-Based Borrowing Constraints and Collateralizing Ideas The second extension is to allow for earnings-based borrowing constraints in the spirit of Lian and Ma (2021) and Greenwald et al. (2019). These empirical studies find that most firms' borrowing decisions are not constrained by the value of their capital assets, as in our baseline model, but rather by the value of their earnings. We could incorporate earnings-based constraints into our framework by changing the borrowing constraint to the form

$$b_{jt+1} \leq \widetilde{\theta}(A_{t+1} z_{jt+1})^{1-\alpha} k_{jt+1}^{\alpha}, \tag{14}$$

where $\widetilde{\theta}$ denotes the tightness of the earnings-based constraint. This specification implies that ideas are partially collateralizable because they affect firms' earnings.

The earnings-based borrowing constraint would amplify the positive externalities of innovation. As in our baseline model, innovation alleviates the equity-issuance constraint $d \geq 0$. But in addition, with the earnings-based constraint, new ideas directly shift out everyone's borrowing constraint (the RHS of (14) is increasing in A_{t+1}). Hence, the planner would have a greater incentive to increase innovation in this extension, especially early in the transition.

The Market for Ideas The third extension we consider is to allow firms to trade ideas between each other. In our baseline model, we abstracted from trade in the "market for ideas" by assuming $i_{jt} \geq 0$. However, financially constrained firms may have an incentive to sell their ideas if the returns to investing in capital are sufficiently high. Empirically, these types of trades seem possible through licensing arrangements, patent sales, or even selling parts of the firm itself.

Given the many frictions rife in these types of trades, we propose introducing a market with search frictions in the spirit of Akcigit, Celik and Greenwood (2016). In this extension, firms would choose between spending some fraction $1 - s_{jt}$ of each time period producing output and the remaining fraction s_{jt} searching in the market for ideas (similar to Lucas and Moll (2014)). If the firm matches with another firms with a higher productivity z_{jt}, then it has the option to buy their productivity at the Nash-bargained price; if the firm instead matches with another firm with lower productivity, then it can sell its own productivity. In order to retain the non-rivalry of ideas, we assume that the selling firms retains its own productivity even after the sale. In that sense, the market for ideas facilitates idea diffusion across firms, as in Lucas and Moll (2014) and Perla and Tonetti (2014).

This extension would enrich our model by providing a third source of firm-level growth: technology adoption. We conjecture that low-productivity firms are more likely to adopt than innovate because their time cost of foregone output while searching is relatively low and their expected return from matching with other (higher-productivity) firms is relatively high. Conversely, high-productivity firms are more likely to sell ideas, especially if they are financially constrained and would like to finance capital investment. In addition, this extension would further endogenize the non-rivalry of ideas through the composition of idea trades that emerges from the market for ideas. While we find these new tradeoffs interesting and potentially important, we have abstracted from them in our baseline analysis for the sake of parsimony.

Revolutionary Entrants and Free Entry The final extension we consider relate to the entry process. In our baseline model, we assume that initial entrants are drawn from a distribution with average productivity and low net worth. This choice is motivated by the actual distributions of new entrants in the data and the typical growth dynamics discussed in Section 2. However, we could also accommodate atypical "revolutionary" firms, like Amazon or Facebook, that enter with extremely high-quality ideas (captured by high levels of idiosyncratic productivity z_{jt}). In this case, the distribution of initial entrants would be a mixture between our typical entrants and the revolutionary entrants. We could discipline the share of revolutionary entrants using the average age of the largest firms in the economy; if revolutionary entrants are more common, then they will be represented in the top of the firm-size distribution and push down its average age.

Related to this discussion, we could also endogenize the mass of new entrants with a free entry condition subject to an entry cost. Of course, the constant flow of new entrants in our calibrated BGP is implicitly supported by some cost of entry in the background.[27] However, if policy interventions raise the expected profits of operating a firm, they will induce net entry that our baseline analysis does not capture. While it would be straightforward to incorporate these entry dynamics, their quantitative magnitudes would depend on relatively ad-hoc assumptions about how policies could move through the distribution of potential entrants. We preferred to abstract from these dynamics with the understanding that our baseline analysis remains a lower bound on the welfare gains from the optimal policy.

9 Conclusion

In this paper, we have developed a quantitative framework to study the allocation of investment and innovation across heterogeneous firms subject to financial frictions. Our model predicts a pecking order of firm growth in which small/young firms grow by accumulating capital but large/old firms grow by innovating and producing new ideas. We confirmed that this prediction is consistent with the behavior of investment and innovation in the micro-data. This pecking order implies that financial frictions lower long-run economic growth by delaying innovation in favor of investment. To the extent that the ideas produced by innovation are non-rival, this private choice misallocates funds away from innovation and toward investment. Instead, the optimal response to the externality requires constrained firms to initially cut their investment expenditures in order to finance higher innovation. We showed that both an innovation subsidy and an investment tax cut can partially achieve this policy goal. Taken together, our results highlight the importance of jointly modeling firms' investment and innovation decisions for understanding economic growth.

[27] In Footnote 11, we described how one could re-interpret the decreasing returns to scale in our baseline model as reflecting constant returns in production plus downward-sloping demand. Alternatively, free entry of new firms can achieve constant returns in the aggregate objects even with decreasing returns for each individual firm.

References

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

Aghion, P., and P. Howitt (1992): "A Model of Growth Through Creative Destruction," *Econometrica: Journal of the Econometric Society*, pp. 323–351.

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

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

Bilal, A., N. Engbom, S. Mongey, and G. L. Violante (2021): "Labor market dynamics when ideas are harder to find," Discussion paper, National Bureau of Economic Research.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2020): "Are ideas getting harder to find?," *American Economic Review*, 110(4), 1104–1144.

Buera, F. J., J. P. Kaboski, and Y. Shin (2011): "Finance and development: A tale of two sectors," *American economic review*, 101(5), 1964–2002.

Caballero, R. J., and E. M. Engel (1999): "Explaining investment dynamics in US manufacturing: a generalized (S, s) approach," *Econometrica*, 67(4), 783–826.

Carroll, C. D. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics letters*, 91(3), 312–320.

Cerra, V., and S. C. Saxena (2008): "Growth dynamics: the myth of economic recovery," *American Economic Review*, 98(1), 439–457.

Clementi, G. L., and B. Palazzo (2016): "Entry, exit, firm dynamics, and aggregate fluctuations," *American Economic Journal: Macroeconomics*, 8(3), 1–41.

Cooley, T. F., and V. Quadrini (2001): "Financial markets and firm dynamics," *American economic review*, 91(5), 1286–1310.

Cooper, R. W., and J. C. Haltiwanger (2006): "On the nature of capital adjustment costs," *The Review of Economic Studies*, 73(3), 611–633.

Dávila, E., and A. Korinek (2018): "Pecuniary externalities in economies with financial frictions," *The Review of Economic Studies*, 85(1), 352–395.

Garcia-Macia, D., C.-T. Hsieh, and P. J. Klenow (2019): "How destructive is innovation?," *Econometrica*, 87(5), 1507–1541.

Geanakoplos, J. D., and H. M. Polemarchakis (1986): "Existence, regularity, and constrained suboptimality of competitive allocations," *Essays in Honor of Kenneth J. Arrow: Volume 3, Uncertainty, Information, and Communication*, 3, 65.

Greenwald, D., et al. (2019): *Firm debt covenants and the macroeconomy: The interest coverage channel*, vol. 1. MIT Sloan School of Management.

Grossman, G. M., and E. Helpman (1991): "Quality ladders in the theory of growth," *The review of economic studies*, 58(1), 43–61.

Hennessy, C. A., and T. M. Whited (2005): "Debt dynamics," *The journal of finance*, 60(3), 1129–1165.

Hopenhayn, H., and R. Rogerson (1993): "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of political Economy*, 101(5), 915–938.

Hopenhayn, H. A. (1992): "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 1127–1150.

Howitt, P., and P. Aghion (1998): "Capital accumulation and innovation as complementary factors in long-run growth," *Journal of Economic Growth*, pp. 111–130.

Hurst, E., and B. W. Pugsley (2011): "What do small businesses do?," Discussion paper, National Bureau of Economic Research.

Itskhoki, O., and B. Moll (2019): "Optimal development policies with financial frictions," *Econometrica*, 87(1), 139–173.

Jones, C. I. (1995): "R & D-based models of economic growth," *Journal of political Economy*, 103(4), 759–784.

Jones, C. I., and J. Kim (2018): "A Schumpeterian model of top income inequality," *Journal of political Economy*, 126(5), 1785–1826.

Kehoe, T. J., and D. K. Levine (1993): "Debt-constrained asset markets," *The Review of Economic Studies*, 60(4), 865–888.

Khan, A., and J. K. Thomas (2008): "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics," *Econometrica*, 76(2), 395–436.

——— (2013): "Credit shocks and aggregate fluctuations in an economy with production heterogeneity," *Journal of Political Economy*, 121(6), 1055–1107.

Klette, T. J., and S. Kortum (2004): "Innovating firms and aggregate innovation," *Journal of political economy*, 112(5), 986–1018.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017): "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 132(2), 665–712.

Lian, C., and Y. Ma (2021): "Anatomy of corporate borrowing constraints," *The Quarterly Journal of Economics*, 136(1), 229–291.

Lorenzoni, G. (2008): "Inefficient credit booms," *The Review of Economic Studies*, 75(3), 809–833.

Lucas, R. E., and B. Moll (2014): "Knowledge growth and the allocation of time," *Journal of Political Economy*, 122(1), 1–51.

Luttmer, E. G. (2007): "Selection, growth, and the size distribution of firms," *The Quarterly Journal of Economics*, 122(3), 1103–1144.

Midrigan, V., and D. Y. Xu (2014): "Finance and misallocation: Evidence from plant-level data," *American economic review*, 104(2), 422–458.

Nuño, G., and B. Moll (2018): "Social optima in economies with heterogeneous agents," *Review of Economic Dynamics*, 28, 150–180.

Ottonello, P., D. J. Perez, and P. Varraso (2022): "Are collateral-constraint models ready for macroprudential policy design?," *Journal of International Economics*, 139, 103650.

Ottonello, P., and T. Winberry (2020): "Financial heterogeneity and the investment channel of monetary policy," *Econometrica*, 88(6), 2473–2502.

Perla, J., and C. Tonetti (2014): "Equilibrium imitation and growth," *Journal of Political Economy*, 122(1), 52–76.

Peters, R. H., and L. A. Taylor (2017): "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 123(2), 251–272.

Romer, P. M. (1990): "Endogenous technological change," *Journal of political Economy*, 98(5, Part 2), S71–S102.

Winberry, T. (2021): "Lumpy investment, business cycles, and stimulus policy," *American Economic Review*, 111(1), 364–396.

Zwick, E., and J. Mahon (2017): "Tax policy and heterogeneous investment behavior," *American Economic Review*, 107(1), 217–248.

A Characterizing Firms' Decision Rules and the BGP

This appendix characterizes the individual firm's decisions and defines a balanced growth path. We proceed in three steps. First, we detrend the problem in order to work with a stationary Bellman equation for which the usual numerical tools apply. Second, we characterize the solution of the detrended problem and show that it results in Proposition 1 in the main text. Finally, we use these results to show that all decisions and macroeconomic aggregates scale with the growth rate g in a balanced growth path.

A.1 Detrending

We will scale the problem by average productivity $Z_t = A_t \int z_{jt} dj = \left(\int z_{jt} dj\right)^{1+a}$. To that end, let $\widetilde{n} = \frac{n}{Z_t}$, $\widetilde{k} = \frac{k}{Z_t}$ denote variables relative to Z_t. The only except is that we will define $\widetilde{z} = \frac{z}{\int z_{jt} dj}$. Divide the Bellman equation (4) by Z_t to get

$$\frac{v_t^{\text{cont}}(z,n)}{Z_t} = \max_{k',i,b'} \frac{n}{Z_t} - \frac{k'}{Z_t} - \frac{A_t z i}{Z_t} + \frac{b'}{Z_t(1+r_t)} + \frac{1}{1+r_t}\mathbb{E}_t\left[\pi_d \frac{n'}{Z_t} + (1-\pi_d)\frac{v_{t+1}^{\text{cont}}(z',n')}{Z_t}\right], \quad (15)$$

where we have expanded $\mathbb{E}_t[v_{t+1}(z',n')] = \pi_d \mathbb{E}_t[n'] + (1-\pi_d)\mathbb{E}_t[v_{t+1}(z',n')]$.

Our goal is to write (15) in terms of the detrended variables and the growth rate $g_t = \frac{Z_{t+1}}{Z_t}$ only. To that end, note that $\frac{k'}{Z_t} = \frac{k'}{Z_{t+1}}\frac{Z_{t+1}}{Z_t} = (1+g_t)\widetilde{k}'$. Similarly, $\frac{b'}{Z_t} = (1+g_t)\widetilde{b}'$. Now multiply and divide the continuation value by $\frac{Z_{t+1}}{Z_{t+1}}$ to get

$$\frac{v_t^{\text{cont}}(z,n)}{Z_t} = \max_{k',i,b'} \widetilde{n} - (1+g_t)\widetilde{k}' - \widetilde{z}i + \frac{(1+g_t)\widetilde{b}'}{(1+r_t)} + \frac{1+g_t}{1+r_t}\mathbb{E}_t\left[\pi_d \widetilde{n}' + (1-\pi_d)\frac{v_{t+1}^{\text{cont}}(z',n')}{Z_{t+1}}\right].$$

Define $\widetilde{v}_t(\widetilde{z},\widetilde{n}) = \frac{v_t^{\text{cont}}(z,n)}{Z_t}$ to arrive at our final detrended Bellman equation:

$$\widetilde{v}_t(\widetilde{z},\widetilde{n}) = \max_{\widetilde{k}',i,\widetilde{b}'} \widetilde{n} - (1+g_t)\widetilde{k}' - \widetilde{z}i + \frac{(1+g_t)\widetilde{b}'}{(1+r_t)} + \frac{1+g_t}{1+r_t}\mathbb{E}_t\left[\pi_d \widetilde{n}' + (1-\pi_d)\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')\right]. \quad (16)$$

Finally, we detrend the constraints and consistency conditions of this problem. Clearly, we have $\widetilde{d} \geq 0$, $\widetilde{b}' \leq \theta\widetilde{k}'$, and $\widetilde{n}' = (\widetilde{z}')^{1-\alpha}(\widetilde{k}')^{\alpha} + (1-\delta)\widetilde{k}' - \widetilde{b}'$. In terms of the law of motion for z, in the event of a successful innovation, we have

$$\log \frac{z}{\int z_{jt+1}dj} = \log \frac{z}{\int z_{jt+1}dj} + \Delta + \varepsilon_{jt+1} = \log \frac{z}{\int z_{jt}dj} \frac{\int z_{jt}dj}{\int z_{jt+1}dj} + \Delta + \varepsilon_{jt+1}$$

which implies

$$\log \widetilde{z}' = \log \frac{\widetilde{z}}{1+\widetilde{g}_t} + \Delta + \varepsilon_{jt+1}$$

where $\widetilde{g}_t = \frac{\int z_{jt+1}dj}{\int z_{jt}dj}$ is the growth rate of firm-specific productivity. Since $Z_t = \left(\int z_{jt}dj\right)^{1+a}$, we have $1 + g_t = (1 + \widetilde{g}_t)^{1+a}$.

A.2 Proof of Proposition 1

Our characterization in Proposition 1 is similar to Khan and Thomas (2013), extended to include the innovation decision. We proceed in three steps. First, we set up the Lagrangian and take the associated first-order conditions. Second, we use those first-order conditions to derive the partition of the state space from the first part of Proposition 1. Finally, we un-detrend those first-order conditions to get the system of equations in the second part of Proposition 1.

A.2.1 Lagrangian

The Lagrangian of the detrended Bellman equation (16) is

$$\mathcal{L} = (1+\lambda_t(\widetilde{z},\widetilde{n})) \left(\widetilde{n} - (1+g_t)\widetilde{k}' - \widetilde{z}i + \frac{(1+g_t)\widetilde{b}'}{(1+r_t)}\right) + (1+g_t)\mu_t(\widetilde{z},\widetilde{n}) \left(\theta\widetilde{k}' - \widetilde{b}'\right) \tag{17}$$
$$+ \chi_t(\widetilde{z},\widetilde{n})i + \frac{1+g_t}{1+r_t}\mathbb{E}_t\left[\pi_d\widetilde{n}' + (1-\pi_d)\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')\right],$$

where $\lambda_t(\widetilde{z},\widetilde{n})$ is the multiplier on the no-equity issuance constraint $\widetilde{d} \geq 0$, $\mu_t(\widetilde{z},\widetilde{n})$ is the multiplier on the collateral constraint $\widetilde{b}' \leq \theta\widetilde{k}'$, and $\chi_t(\widetilde{z},\widetilde{n})$ is the multiplier on the nonnegativity constraint on innovation $i \geq 0$.

The first-order condition for borrowing $\widetilde{b}'$ is

$$(1+\lambda_t(\widetilde{z},\widetilde{n}))\frac{1+g_t}{1+r_t} = (1+g_t)\mu_t(\widetilde{z},\widetilde{n}) - \frac{1+g_t}{1+r_t}\mathbb{E}_t\left[\pi_d\frac{\partial\widetilde{n}'}{\partial\widetilde{b}'} + (1-\pi_d)\frac{\partial\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')}{\partial\widetilde{n}'}\frac{\partial\widetilde{n}'}{\partial\widetilde{b}'}\right].$$

From the envelope condition, we have $\frac{\partial \widetilde{v}_t(\widetilde{z},\widetilde{n})}{\partial \widetilde{n}'} = 1+\lambda_t(\widetilde{z},\widetilde{n})$. Use that together with $\frac{\partial \widetilde{n}'}{\partial \widetilde{b}'} = -1$ to get

$$(1+\lambda_t(\widetilde{z},\widetilde{n}))\frac{1+g_t}{1+r_t} = (1+g_t)\mu_t(\widetilde{z},\widetilde{n}) + \frac{1+g_t}{1+r_t}\mathbb{E}_t\left[\pi_d + (1-\pi_d)(1+\lambda_{t+1}(\widetilde{z}',\widetilde{n}'))\right].$$

Note that $\pi_d + (1-\pi_d)(1+\lambda_{t+1}(\widetilde{z},\widetilde{n})) = 1 + (1-\pi_d)\lambda_{t+1}(\widetilde{z},\widetilde{n})$. Use that fact, multiply by $\frac{1+r_t}{1+g_t}$, and subtract 1 from both sides to finally arrive at

$$\lambda_t(\widetilde{z},\widetilde{n}) = (1+r_t)\mu_t(\widetilde{z},\widetilde{n}) + (1-\pi_d)\mathbb{E}_t\lambda_{t+1}(\widetilde{z}',\widetilde{n}'). \tag{18}$$

Hence, the financial wedge $\lambda_t(\widetilde{z},\widetilde{n})$ is the expected value of current and all future Lagrange multipliers on the collateral constraint $\mu_t(\widetilde{z},\widetilde{n})$, discounted by the exit probability.

The first-order condition for capital accumulation $\widetilde{k}'$ is

$$(1+g_t)(1+\lambda_t(\widetilde{z},\widetilde{n})) = \theta(1+g_t)\mu_t(\widetilde{z},\widetilde{n}) + \frac{1+g_t}{1+r_t}\mathbb{E}_t\left[\pi_d\frac{\partial \widetilde{n}'}{\partial \widetilde{k}'} + (1-\pi_d)\frac{\partial \widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')}{\partial \widetilde{n}'}\frac{\partial \widetilde{n}'}{\partial \widetilde{k}'}\right].$$

Note that $\frac{\partial \widetilde{n}'}{\partial \widetilde{k}'} = MPK(\widetilde{z}',\widetilde{k}') + (1-\delta)$, where $MPK(\widetilde{z}',\widetilde{k}') = \alpha\left(\frac{\widetilde{z}'}{\widetilde{k}'}\right)^{1-\alpha}$ is the marginal product of capital. Using very similar steps to above, the terms in the continuation value can be collected to yield

$$1+\lambda_t(\widetilde{z},\widetilde{n}) = \theta\mu_t(\widetilde{z},\widetilde{n}) + \frac{1}{1+r_t}\mathbb{E}_t\left[\left(MPK(\widetilde{z}',\widetilde{k}') + (1-\delta)\right)\left(\pi_d + (1-\pi_d)\lambda_{t+1}(\widetilde{z}',\widetilde{n}')\right)\right]. \tag{19}$$

The first-order condition for innovation i is

$$(1+\lambda_t(\widetilde{z},\widetilde{n}))\widetilde{z} = \chi_t(\widetilde{z},\widetilde{n}) + \frac{1+g_t}{1+r_t}\frac{\partial}{\partial i}\mathbb{E}_t\left[\pi_d\widetilde{n}' + (1-\pi_d)\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')\right].$$

Consider the term in the continuation value in the case where the firm exits in the next period. We can write this expectation as $\mathbb{E}_t[\widetilde{n}'] = \eta(i)\mathbb{E}^{\varepsilon}[\widetilde{n}'|\text{success}] + (1-\eta(i))\mathbb{E}^{\varepsilon}[\widetilde{n}'|\text{failure}]$ where $\mathbb{E}^{\varepsilon}$ denotes the expectation over the idiosyncratic shocks ε. Hence, we have $\frac{\partial \mathbb{E}_t[\widetilde{n}']}{\partial i} = \eta'(i)\left(E^{\varepsilon}[\widetilde{n}'|\text{success}] - E^{\varepsilon}[\widetilde{n}'|\text{failure}]\right)$. By a similar argument,

$$\frac{\partial \mathbb{E}_t[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')]}{\partial i} = \eta'(i)\left(E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\text{success}] - E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\text{failure}]\right).$$

Putting these all together yields

$$(1+\lambda_t(\widetilde{z},\widetilde{n}))\widetilde{z} \geq \frac{1+g_t}{1+r_t}\eta'(i)\begin{bmatrix}\pi_d\left(E^{\varepsilon}[\widetilde{n}'|\text{success}]-E^{\varepsilon}[\widetilde{n}'|\text{failure}]\right)+\\(1-\pi_d)\left(E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\text{success}]-E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\text{failure}]\right)\end{bmatrix}, \quad (20)$$

with equality if $i > 0$.

To summarize, the firm's optimal decisions are characterized by the first-order conditions (18), (19), and (20) together with the complementarity slackness conditions:

$$\mu_t(\widetilde{z},\widetilde{n})(\theta\widetilde{k}'-\widetilde{b}')=0 \text{ with } \mu_t(\widetilde{z},\widetilde{n})\geq 0, \text{ and}$$
$$\lambda_t(\widetilde{z},\widetilde{n})\widetilde{d}=0 \text{ with } \lambda_t(\widetilde{z},\widetilde{n})\geq 0.$$

A.2.2 Partition of State Space

We now use these first order conditions to derive the partition of the state space in the first part of Proposition 1.

Unconstrained Firms We define a financially unconstrained firm as one for whom the financial wedge $\lambda_t(z,n)=0$. From (18), these firms have zero probability of a binding collateral constraint in the future, so $\mu_{jt+s}=\lambda_{jt+s}=0$ for all $s\geq 0$; that is, being unconstrained is an absorbing state. We will guess and verify that these firms decisions are independent of net worth and are characterized by a set of objects $\widetilde{b}_t'^*(\widetilde{z})$, $\widetilde{k}_t'^*(\widetilde{z})$, $i_t^*(\widetilde{z})$, and $\widetilde{v}_t^*(\widetilde{z})$. We now characterize these objects.

First, because $\lambda_t(\widetilde{z},\widetilde{n})=\mu_t(\widetilde{z},\widetilde{n})=0$, they are indifferent over any combination of b' and d which leaves them financially unconstrained. Following Khan and Thomas (2013), we resolve this indeterminacy by assuming firms accumulate the most debt (or, if $b'<0$, do the least amount of savings) which leaves them financially unconstrained. Khan and Thomas (2013) refer to this policy $b_t'^*(\widetilde{z})$ as the *minimum savings policy.* In order to derive a characterization of it, note that if the firm adopts $b_t'^*(\widetilde{z})$ in period t, then its dividends in the next period $t+1$, conditional on a particular realized state $\widetilde{z}'$, are

$$\widetilde{d}_{t+1}(\widetilde{z}')=(\widetilde{z}')^{1-\alpha}(\widetilde{k}_t'^*(\widetilde{z}))^{\alpha}+(1-\delta)\widetilde{k}_t'^*(\widetilde{z})-\widetilde{b}_t'^*(\widetilde{z})-\widetilde{z}'i_{t+1}^*(\widetilde{z}')-(1+g_{t+1})\widetilde{k}_{t+1}'^*(\widetilde{z}')+\frac{1+g_{t+1}}{1+r_{t+1}}\widetilde{b}_{t+1}'^*(\widetilde{z}')$$

In order to be financially unconstrained, it must be the case that $\widetilde{d}_{t+1}(\widetilde{z}') \geq 0$ for all $\widetilde{z}'$ which have a positive probability. The minimum savings policy $\widetilde{b}_t'^*(\widetilde{z})$ is the largest level of debt which satisfies this constraint with probability one:

$$\widetilde{b}_t'^*(\widetilde{z}) = \min_{\widetilde{z}'} (\widetilde{z}')^{1-\alpha}(\widetilde{k}_t'^*(\widetilde{z}))^\alpha + (1-\delta)\widetilde{k}_t'^*(\widetilde{z}) - \widetilde{z}' i_{t+1}^*(\widetilde{z}') - (1+g_{t+1})\widetilde{k}_{t+1}'^*(\widetilde{z}') + \frac{1+g_{t+1}}{1+r_{t+1}}\widetilde{b}_{t+1}'^*(\widetilde{z}') \quad (21)$$

Note that this policy implies dividends are zero at a minimizer of the RHS of (21) and strictly positive otherwise.

Next, we define $\widetilde{v}_t^*(\widetilde{z})$ to be the value of a firm starting right after they adopt the unconstrained policies:

$$\widetilde{v}_t^*(\widetilde{z}) = -(1+g_t)\widetilde{k}_t'^*(\widetilde{z}) - \widetilde{z} i_t^*(\widetilde{z}) + \frac{(1+g_t)\widetilde{b}_t'^*(\widetilde{z})}{1+r_t} + \frac{1}{1+r_t}\mathbb{E}_t\left[\pi_d \widetilde{n}' + (1-\pi_d)\widetilde{v}_{t+1}^*(\widetilde{z}')\right], \quad (22)$$

where $\widetilde{n}' = (\widetilde{z}')^{1-\alpha}(\widetilde{k}_t'^*(\widetilde{z}))^\alpha + (1-\delta)\widetilde{k}_t'^*(\widetilde{z}) - \widetilde{b}_t'^*(\widetilde{z})$ is independent of $\widetilde{n}$. Since the financial constraints never bind for unconstrained firms, their value function is linearly separable in net worth. Therefore, the total value of a firm who becomes unconstrained in period t is $\widetilde{v}_t(\widetilde{z}, \widetilde{n}) = \widetilde{n} + \widetilde{v}_t^*(\widetilde{z})$.

Given this characterization of the value function, the first-order conditions for capital and innovation (19) and (20) become

$$1 = \frac{1}{1+r_t}\mathbb{E}_t[MPK(\widetilde{z}', \widetilde{k}') + (1-\delta)] \quad (23)$$

$$1 \geq \frac{\eta'(i)}{\widetilde{z}}\frac{1+g_t}{1+r_t}\mathbb{E}_t\begin{bmatrix}\pi_d\left(E^\varepsilon[\widetilde{n}'|\text{success}] - E^\varepsilon[\widetilde{n}'|\text{failure}]\right) + \\ (1-\pi_d)\left(E^\varepsilon[\widetilde{v}_{t+1}^*(\widetilde{z}')|\text{success}] - E^\varepsilon[\widetilde{v}_{t+1}^*(\widetilde{z}')|\text{failure}]\right)\end{bmatrix}. \quad (24)$$

Note that the innovation policy implicitly enters the first-order condition for capital (23) through the expectations operator. Nevertheless, one can verify from (19) and (20) that these policies are independent of current net worth $\widetilde{n}$ given that both $\widetilde{n}'$ and $\widetilde{v}_{t+1}^*(\widetilde{z}')$ are themselves independent of net worth.

Finally, note that if it is feasible to follow these policies, then it will also be optimal because they solve the firm's profit maximization problem with an expanded choice set. In turn, it is feasible to follow these policies if the firm can adopt them without violating the no-equity issuance constraint:

$$\widetilde{n} - (1 + g_t)k_t'^*(\widetilde{z}) - \widetilde{z}i_t^*(\widetilde{z}) + \frac{1 + g_t}{1 + r_t}\widetilde{b}_t'^*(\widetilde{z}) \geq 0. \tag{25}$$

This condition is satisfied if and only if $\widetilde{n} \geq \overline{n}_t(\widetilde{z}) \equiv (1 + g_t)\widetilde{k}_t'^*(\widetilde{z}) + \widetilde{z}i_t^*(\widetilde{z}) - \frac{(1+g_t)\widetilde{b}_t'^*(\widetilde{z})}{1+r_t}$.

Constrained Firms We define financially constrained firms as those for whom $\lambda_t(z, n) > 0$, i.e. there is a positive probability of facing a binding collateral constraint. These firms' decision rules are characterized by the full system of first-order conditions (18), (19), and (20), and therefore depend on net worth. We divide these firms into two cases: (i) *currently constrained* firms currently face a binding collateral constraint, i.e., $\mu_t(\widetilde{z}, \widetilde{n}) > 0$, and (ii) *potentially constrained* firms who do not currently face a binding collateral constraint, i.e., $\mu_t(\widetilde{z}, \widetilde{n}) = 0$.

To derive the threshold $\underline{n}_t(\widetilde{z}, \widetilde{n})$ from the proposition, let $i_t^{\mathrm{p}}(\widetilde{z}, \widetilde{n})$, $\widetilde{k}_t'^{\mathrm{p}}(\widetilde{z}, \widetilde{n})$, and $\widetilde{b}_t'^{\mathrm{p}}(\widetilde{z}, \widetilde{n})$ denote the policy rules of the currently constrained firms. If these choices are feasible, then they are also optimal because they solve a relaxed version of the full problem. The policies are feasible as long as

$$\widetilde{n} \geq \underline{n}_t(\widetilde{z}, \widetilde{n}) \equiv \widetilde{z}i_t(\widetilde{z}, \widetilde{n}) + (1 + g_t)\widetilde{k}_t'^{\mathrm{p}}(\widetilde{z}, \widetilde{n}) - \frac{(1 + g_t)b_t'^{\mathrm{p}}(\widetilde{z}, \widetilde{n})}{1 + r_t}.$$

A.2.3 Un-Detrending the Conditions

We now show that the detrended first-order conditions (18), (19), and (20) derived above imply the conditions (5), (6), and (7) from the main text.

We start with the first-order condition for capital. First note that, from the chain rule,

$$\frac{\partial v_t(z, n)}{\partial n} = Z_t \frac{\partial \widetilde{v}_t(\widetilde{z}, \frac{n}{Z_t})}{\partial n} = \frac{Z_t}{Z_t} \frac{\partial \widetilde{v}_t(\widetilde{z}, \widetilde{n})}{\partial \widetilde{n}} \implies 1 + \lambda_t(z, n) = 1 + \lambda_t(\widetilde{z}, \widetilde{n}),$$

i.e. the financial wedge is the same in the detrended and un-detrended problems. Next, note that

$$MPK_{t+1}(z', k') = \alpha \left(\frac{A_{t+1}z'}{k'} \right)^{1-\alpha} = \alpha \left(\frac{Z_t}{Z_t} \frac{A_{t+1}\widetilde{z}'}{\widetilde{k}'} \right)^{1-\alpha} = MPK(z', k').$$

Hence, the detrended first-order condition (19) directly implies the undetrended first-order condition (5) (where $\mu_t(z,n) = \mu_t(\widetilde{z},\widetilde{n})$ as well).

Next, consider the detrended first-order condition for innovation (20). Plugging in the fact that $1 + g_t = \frac{Z_{t+1}}{Z_t}$ and rearranging gives

$$(1+\lambda_t(z,n))Z_t\widetilde{z} \geq \frac{\eta'(i_t(z,n))}{1+r_t} Z_{t+1}\mathbb{E}_t \begin{bmatrix} \pi_d \left(E^{\varepsilon}[\widetilde{n}'|\text{success}] - E^{\varepsilon}[\widetilde{n}'|\text{failure}]\right) + \\ (1-\pi_d)\left(E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\text{success}] - E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\text{failure}]\right) \end{bmatrix}$$

By definition of the detrended variables, this equation is the same as the un-detrended condition (6) from the main text.

The nonnegativity constraint for dividends (7) follows directly from our detrending of the problem.

A.3 Balanced Growth Path

In this subsection, we characterize a balanced growth path of the model. In order to do so, we must first explicitly write out the law of motion for the distribution of firms. We find it easier to work with the distribution over de-trended state variables, $\Phi_t(\widetilde{z},\widetilde{n})$. Heuristically, its evolution is given by

$$\widetilde{\Phi}_{t+1}(\widetilde{z}',\widetilde{n}') = (1-\pi_d)\int\int\int \begin{pmatrix} \eta(i_t(\widetilde{z},\widetilde{n}))\left[\mathbb{1}\{\widetilde{z}' = \frac{\widetilde{z}e^{\Delta}e^{\varepsilon}}{1+\widetilde{g}_t}\} \times \mathbb{1}\{n'(\frac{\widetilde{z}e^{\Delta}e^{\varepsilon}}{1+\widetilde{g}_t}, k_t'(\widetilde{z},\widetilde{n}), b_t'(\widetilde{z},\widetilde{n}))\}\right] \\ +(1-\eta(i_t(\widetilde{z},\widetilde{n})))\left[\mathbb{1}\{\widetilde{z}' = \frac{\widetilde{z}e^{\varepsilon}}{1+\widetilde{g}_t}\} \times \mathbb{1}\{n'(\frac{\widetilde{z}e^{\varepsilon}}{1+\widetilde{g}_t}, k_t'(\widetilde{z},\widetilde{n}), b_t'(\widetilde{z},\widetilde{n}))\}\right] \end{pmatrix}$$
$$\times\, p(\varepsilon)d\varepsilon\widetilde{\Phi}_t(\widetilde{z},\widetilde{n})d\widetilde{z}d\widetilde{n} + \pi_d\widetilde{\Phi}^0(\widetilde{z},\widetilde{n}), \tag{26}$$

where $\widetilde{n}' = (\widetilde{z}')^{1-\alpha}(\widetilde{k}_t'(\widetilde{z},\widetilde{n}))^{\alpha} + (1-\delta)\widetilde{k}_t'(\widetilde{z},\widetilde{n}) - \widetilde{b}_t'(\widetilde{z},\widetilde{n})$ is the law of motion for detrended state variables induced by the policy rules.[28]

We are now ready to define a **balanced growth path** as the limiting behavior of the model when $\frac{Z_{t+1}}{Z_t} = 1 + g$ for all t. Using the results in the previous subsections, we have shown that the firm value function and decision rules are all scaled by Z_t in the sense that

[28] This description is heuristic because the true transition function for the distribution should be defined over measurable sets of $(\widetilde{z}',\widetilde{n}')$. One can view the heuristic evolution (26) as the generator of that transition function if one interprets the indicator functions $\mathbb{1}$ as Dirac delta functions.

their detrended analogs $\widetilde{v}(\widetilde{z},\widetilde{n})$ are time-invariant. In addition, the distribution of detrended state variables $\widetilde{\Phi}(\widetilde{z},\widetilde{n})$ is constant and equal to the stationary distribution implied by (26). Finally, it is easy to see that aggregate consumption is stationary because can be written as the integral of the policy rules, which scale with Z_t, against the stationary distribution:

$$C = \int \widetilde{z}^{1-\alpha}\widetilde{k}^{\alpha}d\widetilde{\Phi}(\widetilde{z},\widetilde{k},\widetilde{b}) - (1-\pi_d)\int \left(((1+g)\widetilde{k}'_t(\widetilde{z},\widetilde{k},\widetilde{b}) - (1-\delta)\widetilde{k}) + \widetilde{z}i_t(\widetilde{z},\widetilde{k},\widetilde{b})\right) d\widetilde{\Phi}(\widetilde{z},\widetilde{k},\widetilde{b})$$
$$- \pi_d \int \widetilde{k}d\widetilde{\Phi}^0(\widetilde{z},\widetilde{k},\widetilde{b}),$$

where (abusing notation somewhat) $\widetilde{\Phi}(\widetilde{z},\widetilde{k},\widetilde{b})$ denotes the stationary distribution over $(\widetilde{z},\widetilde{k},\widetilde{b})$.

B Solution Algorithm

This appendix describes our numerical solution algorithm. This algorithm may be of interest to other researchers because it is extremely efficient by avoiding numerical optimizer or equation-solver.

Balanced Growth Path We first describe how we solve for a balanced growth path and then describe how we solve for a transition path starting from an arbitrary initial condition away from the BGP. Our algorithm for solving the balanced growth path iterates over candidate growth rates g^*. For each candidate growth rate, the most difficult part is solving for the individual decisions.

We solve for the individual decision rules in two steps. First, we solve for the decisions of the financially unconstrained firms. The key step in this process is iterating over the unconstrained policies $\widetilde{k}'^*_{(\text{it})}(\widetilde{z})$, $i^*_{(\text{it})}(\widetilde{z})$, and $\widetilde{v}_{(\text{it})}(\widetilde{z})$, where (it) indexes the iteration. Given the current iteration of these objects, we perform the following:

(i) Update the investment policy from (19), which becomes $\widetilde{k}'^*_{(\text{it})+1}(\widetilde{z}) = \left(\alpha \frac{\mathbb{E}_t[(\widetilde{z}')^{1-\alpha}]}{r^*-\delta}\right)^{\frac{1}{1-\alpha}}$, where $r^* = \frac{1}{\beta}(1+g^*)^{\sigma} - 1$ is the real interest rate associated with the growth rate g^*. Note that we use the previous iteration of the innovation policy $i^*_{(\text{it})}(\widetilde{z})$ to evaluate the expectation.

(ii) Update the innovation policy from (20), which can also be evaluated in closed form:

$$i^*_{(\text{it})+1}(\widetilde{z}) = \max\{0, \frac{1}{\eta_0}\log\left(\frac{\eta_0}{\widetilde{z}}\frac{1+g^*}{1+r^*}\mathbb{E}_t\left[\begin{matrix}\pi_d\left(E^\varepsilon[\widetilde{n}'|\text{success}] - E^\varepsilon[\widetilde{n}'|\text{failure}]\right) + \\ (1-\pi_d)\left(E^\varepsilon[\widetilde{v}^*_{(\text{it})}(\widetilde{z}')|\text{success}] - E^\varepsilon[\widetilde{v}^*_{(\text{it})}(\widetilde{z}')|\text{failure}]\right)\end{matrix}\right]\right)\}.$$

We use the new iteration of the capital policy $k'^*_{(\text{it})+1}(\widetilde{z})$ to evaluate the evolution of net worth. Note that the minimum savings policy drops out of this difference and is therefore not necessary for this computation.

(iii) Update the value function $\widetilde{v}^*_{(\text{it})+1}(\widetilde{z})$ by iterating on the Bellman operator implied by (22).

Given these unconstrained objects, we can solve for the minimum savings policy by iterating on the operator implied by (21). Finally, we can recover the unconstrained net worth cutoff $\overline{n}(\widetilde{z})$ from (25).

With these unconstrained policies in hand, we can now solve for he decision rules for all firms over the entire state space $(\widetilde{z},\widetilde{n})$. We do so by iterating on $\widetilde{k}'_{(\text{it})}(\widetilde{z},\widetilde{n})$, $\widetilde{b}'_{(\text{it})}(\widetilde{z},\widetilde{n})$, $i_{(\text{it})}(\widetilde{z},\widetilde{n})$, $\lambda_{(\text{it})}(\widetilde{z},\widetilde{n})$, and $v_{(\text{it})}(\widetilde{z},\widetilde{n})$:

(i) If a particular state $(\widetilde{z},\widetilde{n})$ satisfies $\widetilde{n} > \overline{n}(\widetilde{z})$, then use the unconstrained policies and value derived above.

(ii) Solve for the policy rules assuming the collateral constraint is not binding:

- Update the capital accumulation policy from (19), which can be computed in closed form:

$$\widetilde{k}'_{(\text{it})+1}(\widetilde{z},\widetilde{n}) = \left(\alpha\frac{\mathbb{E}_t[(\widetilde{z}'(1+1-\pi_d)\lambda_{(\text{it})}(\widetilde{z}',\widetilde{n}')]}{(1+r^*)(1+\lambda_{(\text{it})}(\widetilde{z},\widetilde{n})) - (1-\delta)\mathbb{E}_t[(1+1-\pi_d)\lambda_{(\text{it})}(\widetilde{z}',\widetilde{n}')]}\right)^{\frac{1}{1-\alpha}},$$

where we compute the law of motion for net worth $\widetilde{n}$ and the expectation using the current iteration (it) of the policy rules.

- Update the implied $\widetilde{b}'_{(\text{it})+1}$ from the $\widetilde{d}=0$ constraint:

$$\widetilde{b}'_{(\text{it})+1}(\widetilde{z},\widetilde{n}) = \frac{1+r^*}{1+g^*}\left(\widetilde{z}i_{(\text{it})}(\widetilde{z},\widetilde{n}) + (1+g^*)\widetilde{k}'_{(\text{it})+1}(\widetilde{z},\widetilde{n}) - \widetilde{n}\right).$$

(iii) For each point in the state space $(\widetilde{z}, \widetilde{n})$, which if the collateral constraint is binding at these candidate solutions, i.e. if $\widetilde{b}'_{(\text{it})+1}(\widetilde{z}, \widetilde{n}) > \theta \widetilde{k}'_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$. If so, compute the policies with a binding collateral constraint:

- Update the capital accumulation policy from the $\widetilde{d} = 0$ constraint with $\widetilde{b}' = \theta \widetilde{k}'$:

$$\widetilde{k}'_{(\text{it})+1} = \frac{\widetilde{n} - \widetilde{z} i_{(\text{it})}(\widetilde{z}, \widetilde{n})}{(1+g^*)(1 - \frac{\theta}{1+r^*})}.$$

- Set $\widetilde{b}'_{(\text{it})+1}(\widetilde{z}, \widetilde{n}) = \theta \widetilde{k}'_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$.
- Recover the Langrange multiplier on the collateral constraint $\mu_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$ from the capital Euler equation (19).

(iv) Update the innovation policy (20) given this new iteration of the investment and borrowing policies:

$$i^*_{(\text{it})+1}(\widetilde{z}) = \max\{0, \frac{1}{\eta_0} \log \left(\frac{\eta_0}{\widetilde{z}} \frac{1+g^*}{1+r^*} \mathbb{E}_t \begin{bmatrix} \pi_d \left(E^\varepsilon[\widetilde{n}'|\text{success}] - E^\varepsilon[\widetilde{n}'|\text{failure}] \right) + \\ (1-\pi_d) \left(E^\varepsilon[\widetilde{v}^*_{(\text{it})}(\widetilde{z}')|\text{success}] - E^\varepsilon[\widetilde{v}^*_{(\text{it})}(\widetilde{z}')|\text{failure}] \right) \end{bmatrix} \right) \}$$

where we evaluate the law of motion for net worth using $\widetilde{k}'_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$ and $\widetilde{b}'_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$.

(v) Update the value function $\widetilde{v}_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$ by iterating on the Bellman operator from (16).

(vi) Update the financial wedge $\lambda_{(\text{it})+1}(\widetilde{z}, \widetilde{n})$ from (18):

$$\lambda_{(\text{it})+1}(\widetilde{z}, \widetilde{n}) = (1+r^*)\mu_{(\text{it})+1}(\widetilde{z}, \widetilde{n}) + (1-\pi_d)\mathbb{E}_t[\lambda_{(\text{it})}(\widetilde{z}', \widetilde{n}')].$$

While we do not have a formal proof that this iteration will converge, we find that it robustly converges for the parameterizations that we have explored. Given these policy rules, we compute the stationary distribution $\widetilde{\Phi}(\widetilde{z}, \widetilde{n})$ implied by (26). We now need to compute the aggregate growth rate implied by these decision rules. By definition, the growth rate is $1 + \widehat{g} = (1+g_z)^{1+a}$, where g_z is the growth rate of firm-level productivity z. We compute g_z

using the definition

$$1 + g_z = \frac{(1-\pi_d)\int z' p(\varepsilon)\Phi(s) d\varepsilon ds + \pi_d(1+g_z)\int z\Phi(s)ds}{\int z\Phi(s)ds}$$

where $s = (z, n)$ denotes the individual state vector and $Phi(s)$ is the p.d.f. of incumbent firms. The second term in the numerator reflects our assumption that the average productivity of initial entrants is equal to the average productivity of incumbents. Rearranging this expression gives

$$1 + g_z = \frac{\int z' p(\varepsilon)\Phi(s) d\varepsilon ds}{\int z\Phi(s)ds}.$$

The numerator in this integral is

$$\int \left[\eta(i(s))e^{\Delta}e^{\varepsilon}z + (1-\eta(i(s)))e^{\varepsilon}z\right] p(\varepsilon)\Phi(s)d\varepsilon ds$$
$$= e^{\sigma_\varepsilon^2/2}\left[\int z\Phi(s)ds + \int \eta(i(s))(e^{\Delta}-1)z\Phi(s)ds\right]$$

where the second line uses the fact that ε is log-normally distributed independent of s. Collecting terms, we have

$$1 + g_z = e^{\sigma_\varepsilon^2/2}\left[1 + (e^{\Delta}-1)\frac{\int \eta(i(s))z\Phi(s)ds}{\int z\Phi(s)ds}\right] \quad \Longrightarrow \; 1 + \widehat{g} = (1+g_z)^{1+a_0}.$$

The above procedure defines a mapping from the current guess of the growth rate, g^*, to a new guess $\widehat{g} = f(g^*)$. A balanced growth path is a fixed point of this mapping. We compute that fixed point using the bisection method.

Transition Path We can solve for the transition path starting at an arbitrary initial distribution $\widetilde{\Phi}_0(\widetilde{z}, \widetilde{n})$ using a nonlinear equation solver. Specifically, we assume the economy converges to the balanced growth path by some finite period T and define the transition path as a sequence of $\{g_t, r_t\}_{t=0}^T$ which solves $f(\{g_t, r_t\}) = 0$, where f performs the following:

(i) Given the sequence $\{g_t, r_t\}_{t=0}^T$, solve for the individual decisions using backward iteration in the scheme described above for computing the BGP.

(ii) Given these policies and the initial distribution, $\widetilde{\Phi}_0(\widetilde{z}, \widetilde{n})$, simulate forward to get the path of distributions $\{\widetilde{\Phi}_t(\widetilde{z}, \widetilde{n})\}_{t=1}^T$.

(iii) The elements of $f(\{g_t, r_t\})$ are then the aggregate consistence conditions:

$$(e^{\Delta} - 1) \int \eta(i_t(\widetilde{z}, \widetilde{n})) d\widetilde{\Phi}_t(\widetilde{z}, \widetilde{n}) - g_t = 0$$
$$\frac{1}{\beta} \left(\frac{C_{t+1}}{C_t} \right)^{\sigma} - (1 + r_t) = 0.$$

C Data Appendix

This appendix provides additional details about the Compustat sample used to calibrate the model in Section 5.

C.1 Data Construction

This subsection describes the firm-level variables used in the empirical analysis of the paper, based on annual Compustat data.

Variables We define the variables used in our empirical analysis as follows:

1. *Investment rate*: ratio of capital expenditures (`capx`) to lagged plant, property, and equipment (`ppegt`).
2. *R&D-to-sales*: ratio of research and development expense (`xrd`) to the average of sales (`sale`) in the previous 5 years.
3. *Cash flows*: measured as the sum of EBITDA and research and development expense divided by lagged plant, property, and equipment.
4. *Capital-to-employment*: defined at the ratio of plant, property, and equipment to employment (`emp`).
5. *Leverage*: defined as the ratio of total debt (sum of `dlc` and `dltt`) to total assets (`at`).

Sample Selection Our empirical analysis excludes:

1. Firms in finance, insurance, and real estate sectors ($\texttt{sic} \in [6000, 6799]$), utilities ($\texttt{sic} \in [4900, 4999]$), nonoperating establishments ($\texttt{sic} = 9995$), and industrial conglomerates ($\texttt{sic} = 9997$).
2. Firms not incorporated in the United States.
3. Firm-year observations that satisfy one of the following conditions, aimed at excluding extreme observations:

   i. Negative assets, sales, capital expenditure, or R&D.

   ii. Low capital values (gross plant, property, and equipment below $5M in 1990 dollars).

   iii. Acquisitions larger than 20% of assets.

   iv. Investment rates higher than 1.

   v. Innovation-to-sales ratios higher than 0.3.

   vi. Leverage higher than 10 or negative.

C.2 Additional Figures and Tables

This subsection collects additional empirical results referenced in the main text.

C.2.1 Descriptive statistics

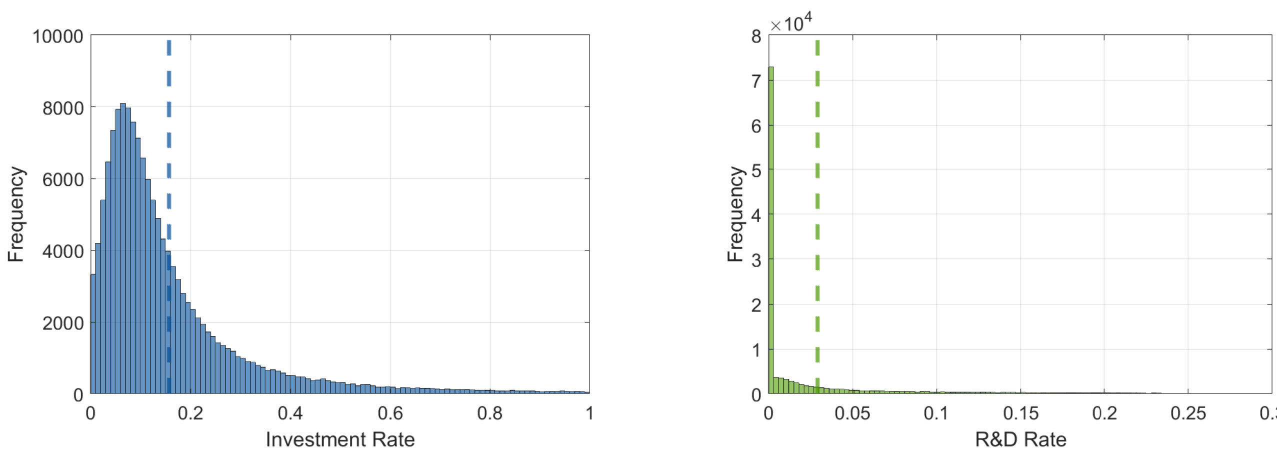
Table 6
Descriptive statistics

	Mean	Median	St dev	95th	Observations
Investment rate	.14	.102	.131	.397	123,954
Investment spike	.198		.4		123,954
Investment rate | spike	.315	.278	.181	.694	23,729
Time since last spike	4.70	2	6.36	18	81,257
R&D-to-sales ratio	.029	0	.057	.170	123,954
Positive R&D expenditure	.446		.497		123,954
R&D-to-sales ratio | positive R&D expenditure	.064	.0345	.070	.223	55,388
Leverage	.258	.229	.215	.659	123,954

Notes: This table shows descriptive statistics for variables used in the empirical analysis of Section 5.1. Investment rate, R&D-to-sales ratio, and leverage are defined in Appendix C.1. *Investment spike* denotes a dummy variable that takes the value of one in periods in which a firm's investment rate is above 20%. *Time since last spike* denotes the number of years since the firm experienced the previous investment spike. *Positive R&D expenditure* denotes a dummy variable that takes the value of one in a period in which a firm's research and development expense (`xrd`) is positive. *Investment rate | spike* and *R&D-to-sales ratio | positive R&D expenditure* report, respectively, moments for investment rates conditional on periods of investment spikes and of R&D-to-sales ratios conditional on positive R&D expenditure. For sample selection, see Appendix C.1.

Figure 8: Distribution of Investment Rates and R&D

(a) Investment rates (b) R&D-to-sales

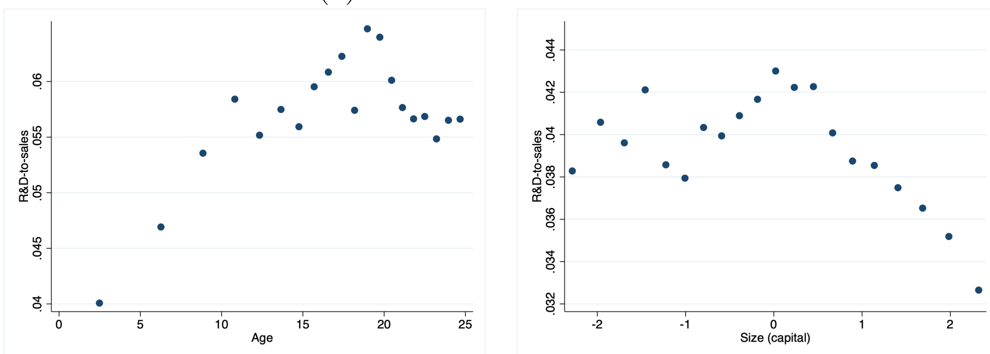


Notes: This figure shows the histogram of investment rates and the R&D-to-sales ratio. Vertical dashed lines represent each variable mean. For variables definitions and sample selection, see Appendix C.1.

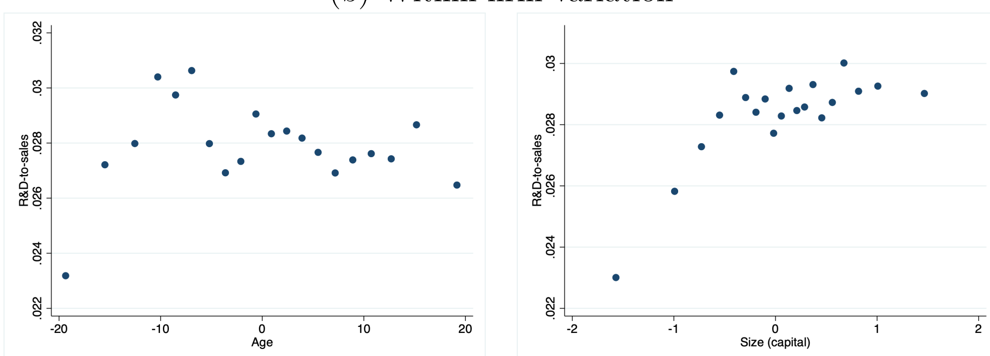
C.2.2 Innovation and Investment across The Size and Age Distribution

FIGURE 9: R&D-to-sales Ratio by Firms' Size and Age: Firm-years with Positive R&D

(a) Within-sector variation



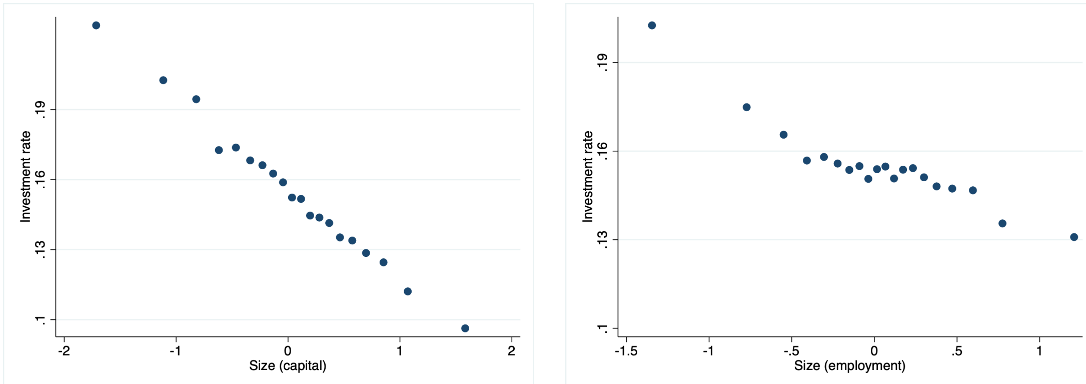
(b) Within-firm variation



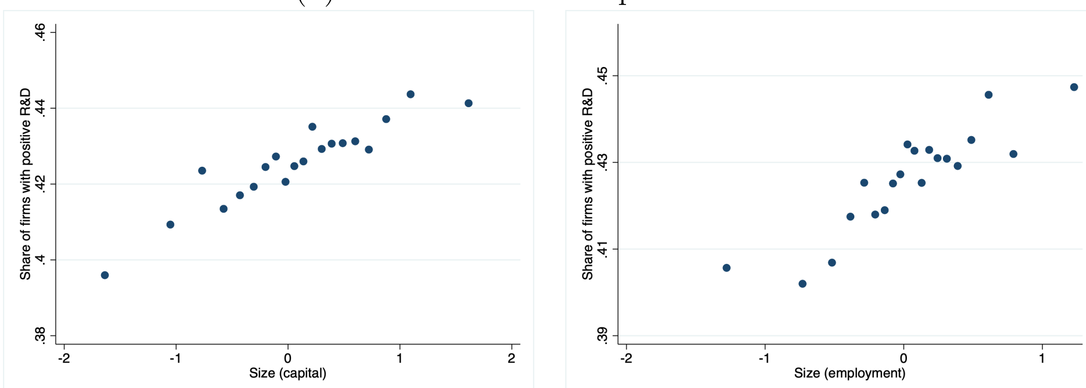
Notes: These figures report binned scatter plots of the R&D-to-sales ratio (for firm-years with positive R&D) by firms' size (measured by the log of real capital) and age of incorporation. Panel (a) reports variables demeaned at 4-digit-NAICS sector level; Panel (b) reports variables demeaned at the firm level, for the sample including firm spells with more than 20 years of data. To construct the plots we add the unconditional mean of R&D-to-sales to demeaned variables. For plots involving age, we restrict observations to firm-years with ages between 0 and 25 years. For variable definitions and sample selection, see Appendix C.1.

Figure 10: Investment Rates and Innovation by Firms' Size: Within-Firm Variation

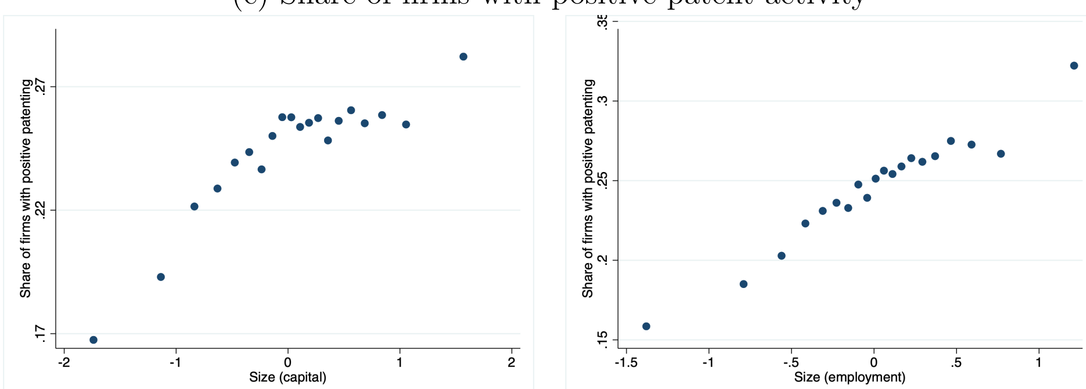
(a) Investment rates



(b) Share of firms with positive R&D



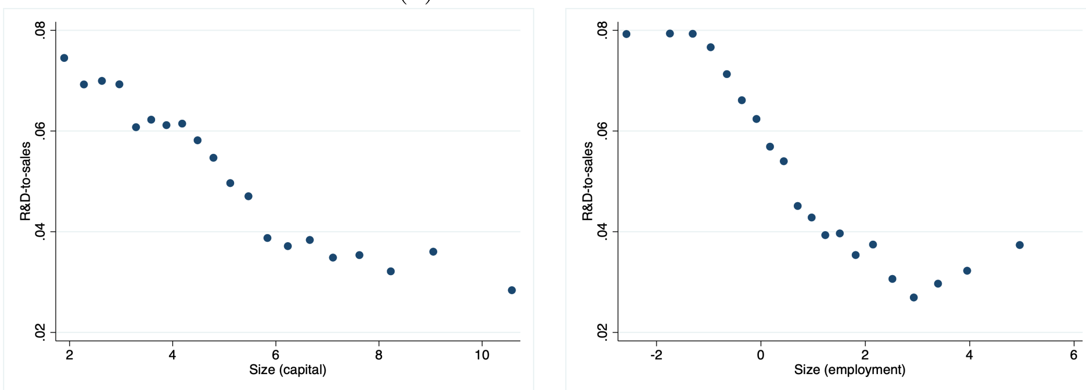
(c) Share of firms with positive patent activity



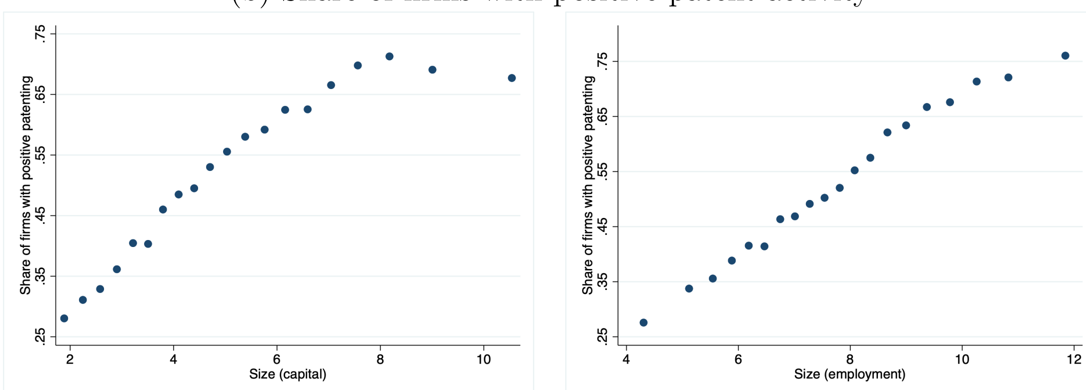
Notes: These figures report binned scatter plots of investment rates, the share of firms with positive R&D, and the share of firms with positive patenting by firms' size, measured by the log of real capital or employment. All variables are demeaned at the firm level, and for each variable, the sample includes only firm spells with more than 20 years of data. To construct the plots for investment rates, the share of firms with positive R&D rates, and the share of firms with positive patenting, we add the unconditional mean of each variable to sector-level demeaned variables. For variable definitions and sample selection, see Appendix C.1.

FIGURE 11: Innovation by Firms' Size: Raw Data for Sample of "Continuously-Innovative Firms"

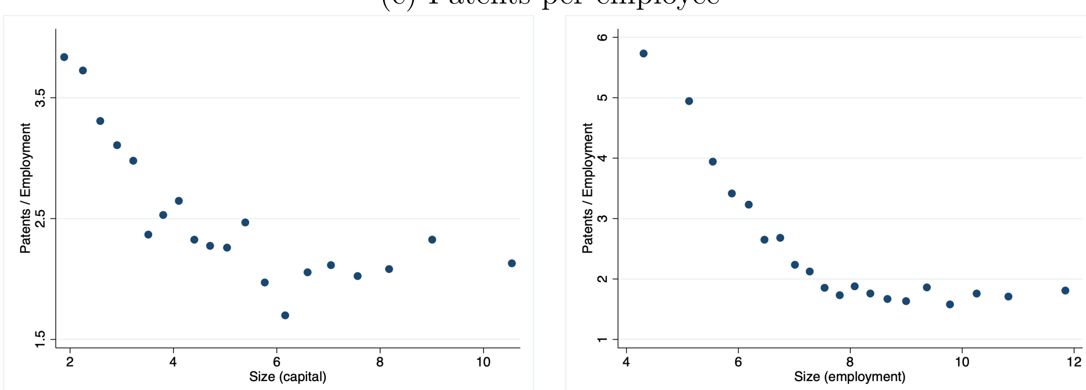
(a) R&D-to-sales ratio



(b) Share of firms with positive patent activity



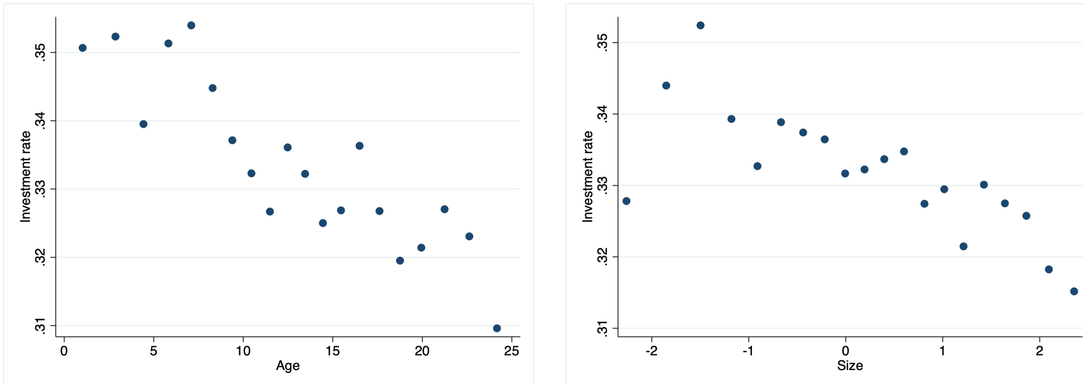
(c) Patents per employee



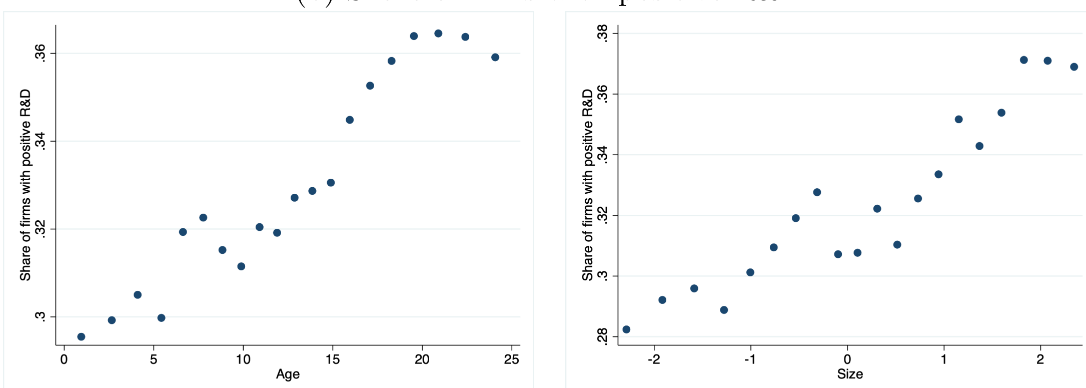
Notes: These figures report binned scatter plots of the R&D-to-sales ratio, the share of firms with positive patenting, and the ratio of patents-to-employees by firms' size, measured by the log of real capital or employment. We report statistics for the sample of "continuously-innovative firms," defined as, those that in the previous five years conducted some positive R&D expenditure and patenting). For variable definitions and sample selection, see Appendix C.1.

Figure 12: Investment Rates and Innovation by Firms' Size and Age: Orbis data

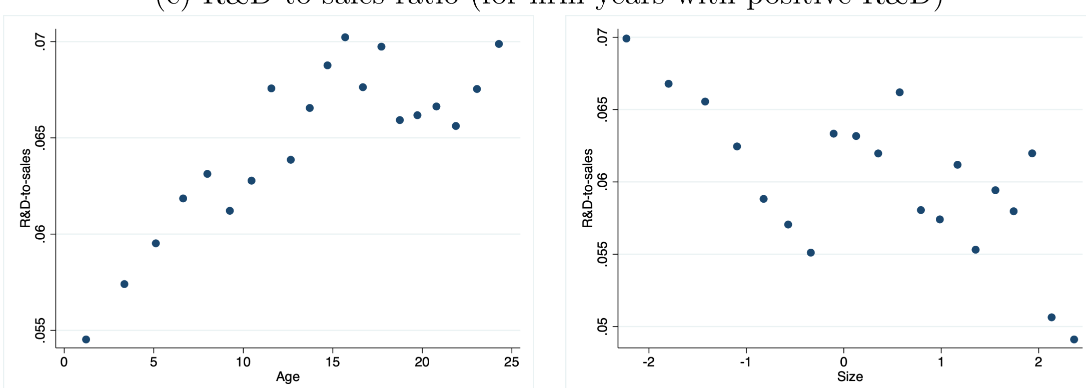
(a) Investment rates



(b) Share of firms with positive R&D



(c) R&D-to-sales ratio (for firm-years with positive R&D)



Notes: These figures report binned scatter plots of investment rates, the share of firms with positive R&D, and the R&D-to-sales ratio (for firm-years with positive R&D) by firms' size (measured by the log of real capital) and age of incorporation, in the Orbis dataset. All variables are demeaned at 4-digit-NAICS sector level. To construct the plots for investment rates, the share of firms with positive R&D rates, the R&D-to-sales ratio, and age variables, we add the unconditional mean of each variable to sector-level demeaned variables. For plots involving age, we restrict observations to firm-years with ages between 0 and 25 years.

C.2.3 Innovation and Investment Spikes

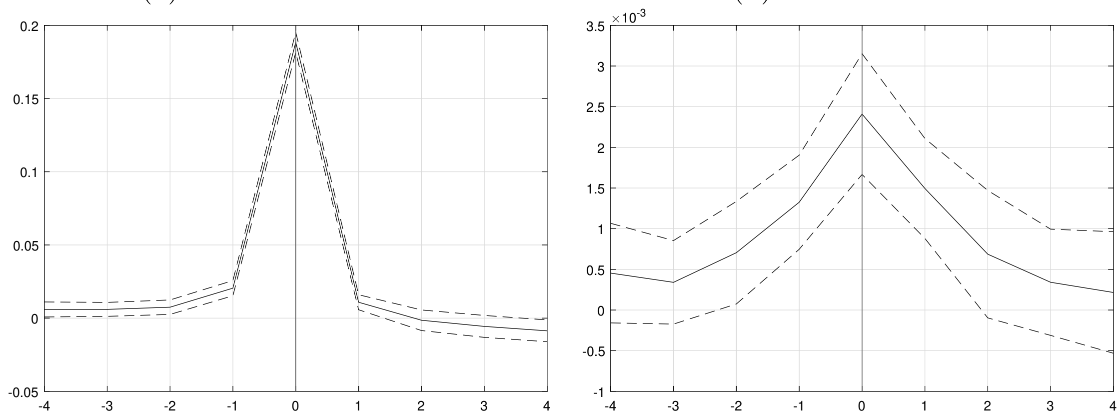
Table 7
Investment spikes and Innovation: Robustness

	(1)	(2)	(3)	(4)	(5)
$\frac{i_{jt-1}}{\tilde{y}_{jt-1}}$	1.12	0.67	1.03	1.12	1.06
	(0.15)	(0.13)	(0.14)	(0.16)	(0.16)
$\frac{\text{cf}_{jt}}{k_{jt}}$	0.11	0.07	0.11	0.12	0.12
	(0.02)	(0.01)	(0.02)	(0.02)	(0.04)
years since spike$_{t-1}$	0.003	0.002	0.003	0.003	0.003
	(0.0008)	(0.0005)	(0.0008)	(0.0008)	(0.0008)
$\frac{k_{jt}}{n_{jt-1}}$	-0.016	-0.01	-.0139	-0.017	-0.014
	(0.003)	(0.002)	(0.003)	(0.004)	(0.003)
Measure of spikes	Absolute	Sectoral	Absolute	Absolute	Absolute
Lags	4	4	3	5	4
Additional controls	No	No	No	No	Size, sales growth, current assets
Observations	53,577	53,577	58,066	49,116	52,292
Adj. R^2	0.282	0.186	0.285	0.279	0.30

Notes: Results from estimating alternative versions of
$\mathbb{1}\{\frac{x_{jt}}{k_{jt}} \geq \chi_s\} = \alpha_j + \alpha_{st} + \sum_{h=1}^{H} \beta_h \left(\frac{\widehat{i}_{jt-h}}{\tilde{y}_{jt-h}}\right) + \Gamma' X_{jt} + \varepsilon_{jt}$, where $\frac{x_{jt}}{k_{jt}}$ denotes the investment rate of firm j in period t; χ_s is a threshold defining investment spikes; $\frac{\widehat{i}_{jt}}{\tilde{y}_{jt}}$ the R&D-to-sales ratio; α_j and α_{st} firm and time by sector fixed effects; $X_{j,t}$ is a vector of firm-level controls; and ε_{jt} is a random error term. Column (1) reports estimates for the baseline specification of Table 1, with $\chi_s = 0.2$, $H = 1$, and the vector X_{jt} including cash flows ($\frac{\text{cf}_{jt}}{k_{jt}}$) and the lumpy-investment controls (years since the last investment spike, years since spike$_{t-1}$, and the standardized capital-output ratio, $\frac{k_{jt}}{n_{jt-1}}$). Column (2) uses a "sectoral" threshold for investment spikes, where $\chi_{ts} = 0.2$ is the mean plus one standard deviation of the distribution of investment rates of sector s (at 2-digit NAICS level). Columns (3) and (4) report results for alternative lags of the R&D-to-sales ratio: $H = 3$ and $H = 5$. Column (5) includes additional control variables: size (measured with the log of real plant, property, and equipment), sales growth, and the share of current assets. For variable definitions and descriptive statistics, see Appendix C.

FIGURE 13: Event-Time Analysis

(a) Investment rates (b) R&D-to-sales



Notes: This figure shows the dynamics of investment rates and R&D-to-sales around investment spike episodes. The figure reports the coefficients β_h from estimating $y_{jt} = \alpha_j + \alpha_{st} + \sum_{h=-4}^{4} \beta_h \frac{x_{jt}}{k_{jt-g}} \geq 0.2 + \Gamma' X_{j,t-1} + \varepsilon_{jt}$, where y_{jt} denotes the investment rate $(\frac{x_{jt}}{k_{jt}})$ or R&D-to-sales ratio $(\frac{\widehat{i}_t}{\widehat{y}_t})$; α_j and α_{ts} firm and time by sector fixed effects; and ε_{jt} is a random error term. For variable definitions and descriptive statistics, see Appendix.

D Details on Calibrated Model

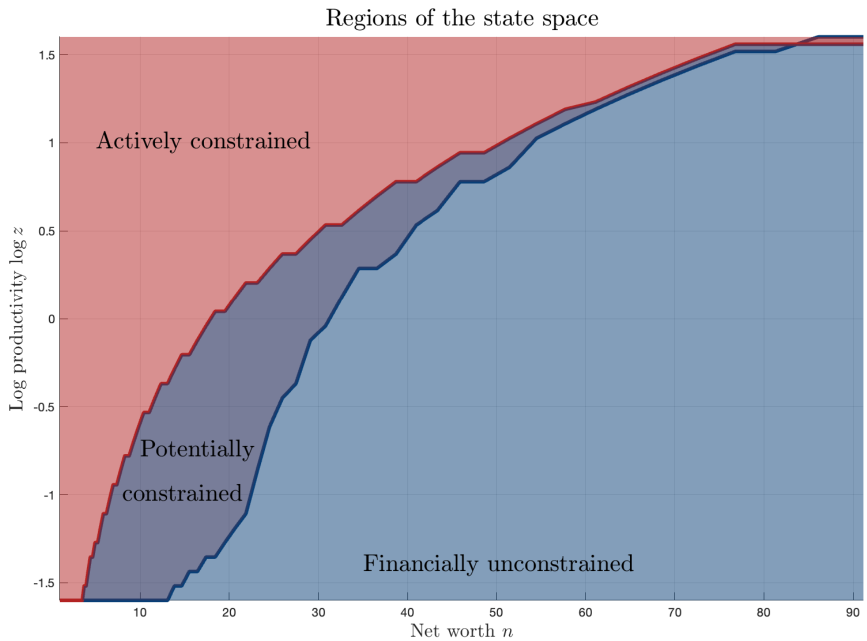
This appendix studies firms' decisions in the calibrated balanced growth path and confirms the model matches various untargeted moments in the data.

D.1 Sources of Firm Heterogeneity

Figure 14 visualizes the partition of the state space characterized in Proposition 1. The red isocurve implicitly defines the constrained cutoff $\underline{n}(z, n)$; firms above this curve are actively constrained. The level of net worth below which firms are constrained is increasing in productivity z because higher productivity firms have a higher optimal scale of capital $k^*(z)$ and therefore a greater incentive to borrow. The blue isocurve implicitly defines the unconstrained cutoff $\overline{n}(z)$; firms below this curve are financially unconstrained. Firms in between these two isocurves are potentially unconstrained.

Decision Rules Figure 15 plots firms' value functions and decision rules as a function of net worth n for different levels of productivity z. Consistent with the pecking order of

FIGURE 14: Partition of the State Space

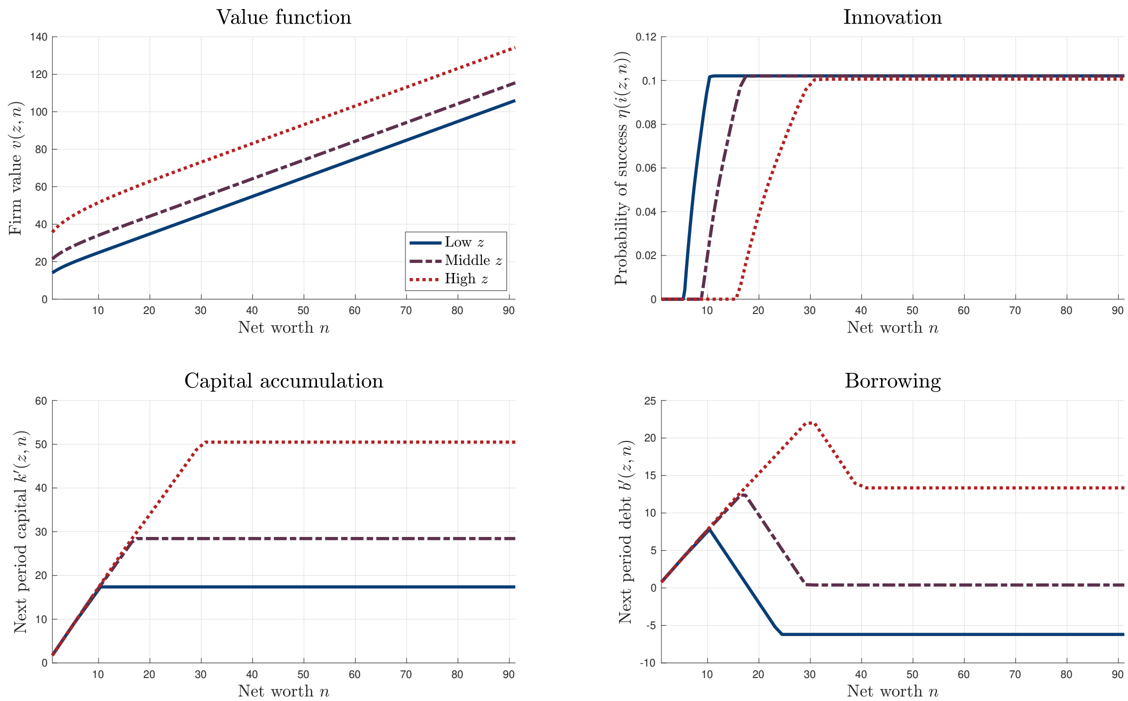


Notes: partition of the state space from Proposition 1 in the market BGP. Net worth n and log productivity $\log z$ have been detrended following Appendix A.

firm growth from Section 4, firms with low net worth spend all their available resources on investment and do not innovate. The level of net worth at which firms begin innovating is increasing in their productivity because higher-productivity firms have a higher marginal product of capital and, therefore, a higher opportunity cost of innovations. While constrained, firms accumulate debt until they reach their optimal scale $k^*(z)$, at which point they use additional net worth to pay down their debt (and potentially engage in financial saving). Once firms become financially unconstrained, they adopt the minimum savings policy described in Proposition 1. Unconstrained firms' capital varies substantially, but all unconstrained firms have the same innovation rate because the cost of innovation is scaled by productivity.

Figure 16 plots the "cash flow sensitivities" of investment and innovation, defined as $\frac{\partial k'(z,n)}{\partial n}$ and $\frac{\partial i(z,n)}{\partial n}$. Of course, unconstrained firms have sensitivities of zero because their decision rules are independent of net worth (see Figure 15). Among constrained firms,

FIGURE 15: Decision Rules



Notes: firm decision rules in the market BGP. All variables have been detrended following Appendix A.

those that do not innovate simply put all additional net worth toward investment. We can explicitly compute the resulting investment-cash flow sensitivity by differentiating the flow of funds constraint (7) with innovation $i(z,n) = 0$ and borrowing $b' = \theta k'$

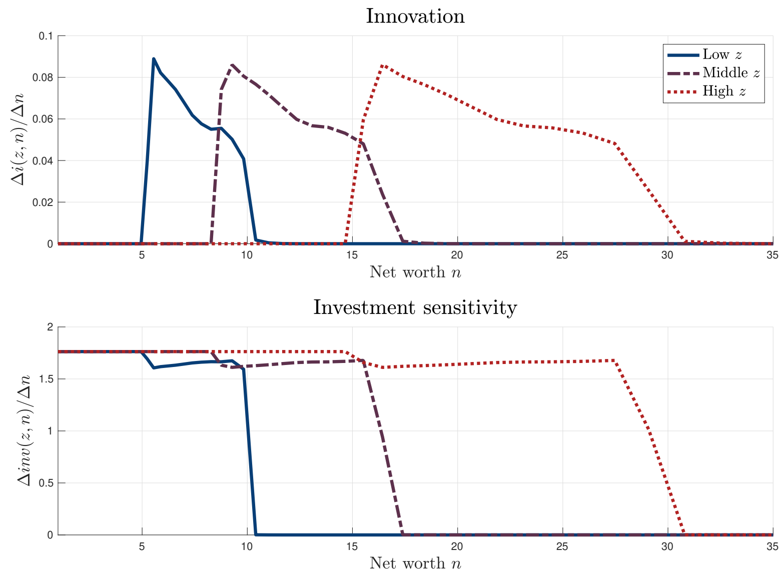
$$k'(z,n) = n + \frac{\theta k'(z,n)}{1+r} \implies \frac{\partial k'(z,n)}{\partial n} = \left(1 - \frac{\theta}{1+r}\right)^{-1} \approx 1.75,$$

where the last approximation uses our calibrated values of $\theta = 0.45$ and $r = 0.04$. Since firms can lever up investment with borrowing, their investment-cash flow sensitivities are above one. Constrained firms with positive innovation have a smaller investment-cash flow sensitivity because they put some of the additional funds toward innovation as well:

$$k'(z,n) + A_t z i(z,n) = n + \frac{\theta k'(z,n)}{1+r} \implies \frac{\partial k'(z,n)}{\partial n} = \left(1 - \frac{\theta}{1+r}\right)^{-1} \left(1 - A_t z \frac{\partial i(z,n)}{\partial n}\right).$$

Quantitatively, Figure 16 shows that the innovation-cash flow sensitivities are an order of magnitude smaller than the investment-cash flow sensitivities.

FIGURE 16: Cash Flow Sensitivities



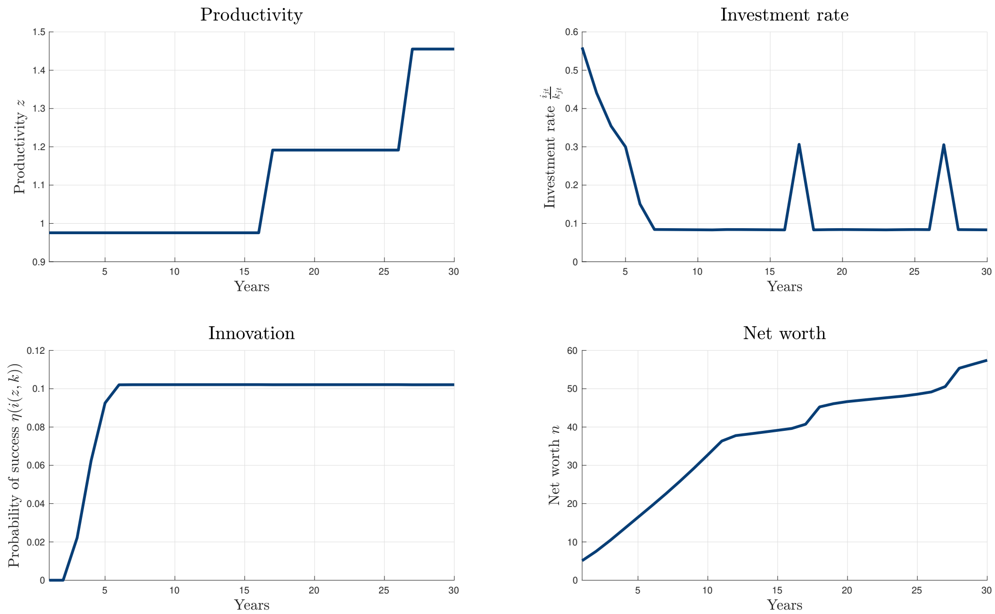
Notes: cash flow sensitivities computed as $\frac{\partial k'(z,n)}{\partial n}$ and $\frac{\partial i(z,n)}{\partial n}$. Derivatives computed using finite differences.

Lifecycle Dynamics Figure 17 plots a sample lifecycle for a firm that enters the economy at time $t = 0$. In order to highlight the role of innovation, we assume that the firm receives no idiosyncratic productivity shocks $\varepsilon_{jt} = 0$ over this sample path. In its first few years of life, the firm has a very high investment rate and does not innovate. But as the firm ages, it exhausts its marginal product of capital, reducing its investment rate and increasing its innovation rate. These dynamics are consistent with the descriptive evidence from Figure 9 in the main text. In this particular sample path, the firm receives two successful innovations: one in year 17 and the other in year 27. Both of these successful innovations are accompanied by investment spikes.

D.2 Distribution of Investment, Innovation, and Leverage

Table 8 compares a number of moments of the stationary distribution of investment, R&D, and leverage from our model to their counterparts in the Compustat data. The model endogenously matches the average investment rate and unconditional frequency of investment spikes fairly well even though they were not directly targeted in the calibration. The model's

FIGURE 17: Sample Firm Lifecycle



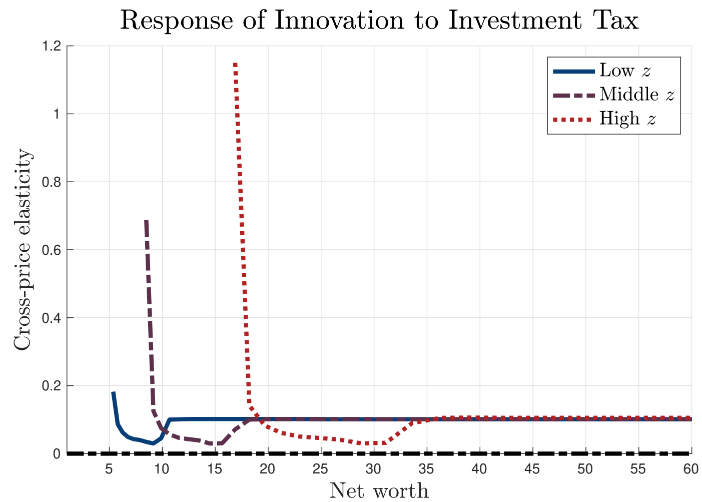
Notes: sample lifecycle profile for a firm without idisosyncratic shocks $\varepsilon_{jt} = 0$ for all j. Initially endowed with approximately average productivity and net worth among new entrants .

TABLE 8
DISTRIBUTION OF INVESTMENT, INNOVATION, AND LEVERAGE

Statistic	**Data**	**Model**
Investment spending		
$\sigma(x_{jt}/k_{jt})$ (targeted)	0.13	0.10
$\mathbb{E}[x_{jt}/k_{jt}\|\text{spike}]$ (targeted)	0.32	0.32
$\mathbb{E}[x_{jt}/k_{jt}]$	0.14	0.14
Frac$(x_{jt}/k_{jt} > 0.2)$	0.21	0.24
R&D spending		
$\mathbb{E}[i_{jt}/y_{jt}]$	0.03	0.07
$\sigma(i_{jt}/y_{jt})$	0.06	0.02
Leverage		
$\mathbb{E}[b_{jt}/k_{jt}]$ (targeted)	0.15	0.13
$\sigma(b_{jt}/k_{jt})$	0.30	0.24
$\mathbb{E}[b_{jt}/k_{jt}]$ (gross)	0.26	0.19
$\sigma(b_{jt}/k_{jt})$ (gross)	0.22	0.18

Notes: cross-sectional statistics from stationary distribution of firms. As in the maint text, x_{jt} denotes investment, k_{jt} denotes capital, i_{jt} denotes innovation, y_{jt} denotes sales, and b_{jt} denotes borrowing. We compute gross borrowing in the model as $\max\{b_{jt}, 0\}$.

FIGURE 18: Heterogeneous Responses to the Bonus Depreciation Allowance



Notes: cross-price elasticity $\frac{\partial \log i(z,n)}{\partial \log(1-\zeta_t)}$ in response to a shock equivalent to a 50% bonus depreciation allowance which reverts back to its long-run average following an AR(1) process with annual persistence 0.75 (giving a half-life of roughly two years). Elasticities computed in impact period of the shock.

average R&D-to-sales ratio is about twice as high as in the data. We chose not to target R&D spending because it is well-known to under-report total innovation expenditures; for example, about half of our firm-year observation report zero R&D expenditures. Conditional on recording R&D spending, the average R&D-to-sales ratio is 0.064 in the data vs. 0.065 in our model. Finally, our model captures most of the dispersion in leverage and also fits the first two moments of the gross leverage distribution fairly well.

D.3 Cross-Price Elasticity of Innovation w.r.t. Investment

Figure 18 shows that the cross-price elasticity of innovation w.r.t. investment is higher for constrained firms than unconstrained firms, as referenced in the main text.

E Constrained Efficiency (Proof of Proposition 4)

We formulate the planner's problem recursively. For notational convenience, let $s = (z, k, b)$ denote a firm type. The planner's state variable is the distribution of firms, $\Phi(s)$. The planner's value function solves the Bellman equation

$$W_t(\Phi) = \max_{k'(\cdot),i(\cdot),b'(\cdot)} \frac{C^{1-\sigma}-1}{1-\sigma} + \beta W_{t+1}(T(\Phi;k'(\cdot),i(\cdot),b'(\cdot))) \text{ such that} \tag{27}$$

$$C = \int \left[(Az)^{1-\alpha}k^{\alpha} + (1-\delta)k\right]\Phi(s)ds - (1-\pi_d)\int \left[k'(s) + Azi(s)\right]\Phi(s)ds \tag{28}$$
$$- \pi_d \int k'\Phi^0(z',k',b')dz'dk'db'$$

$$(Az)^{1-\alpha}k^{\alpha} + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1+r_t} \geq 0 \text{ for all } s \tag{29}$$

$$b'(s) \leq \theta k'(s) \text{ for all } s \tag{30}$$

$$A = \left(\int z\Phi(s)dz\right)^a \tag{31}$$

$$T(\Phi;k'(\cdot),i(\cdot),b'(\cdot))(z',k',b') = \pi_d\Phi^0(z',k',b') \tag{32}$$
$$+ (1-\pi_d)\int \begin{bmatrix} (\mathbb{1}\{k'=k'(s)\}\times \mathbb{1}\{b'=b'(s)\})\times \\ (\eta(i(s))\mathbb{1}\{z'=ze^{\Delta}e^{\varepsilon}\} + (1-\eta(i(s)))\mathbb{1}\{z'=ze^{\varepsilon}\}) \end{bmatrix} p(\varepsilon)\Phi(s)ds,$$

where $p(\varepsilon)$ is the p.d.f. of ε and $T(\Phi;k'(\cdot),i(\cdot),b'(\cdot))$ is the transition function for the distribution. We denote the entire decision rule function using, e.g., $k'(\cdot)$, and the function evaluated at a particular using $k'(s)$.

The planner's problem (27) is a *functional equation* because both the state variable and choice variables are functions of the individual state s. Nuño and Moll (2018) provide conditions under which Lagrangian methods apply using Gâteaux derivatives, which we assume hold in our model as well. These derivatives are the natural extension of partial derivatives into the function space. For example, $\frac{\delta W}{\delta\Phi(s)}(\Phi)$ denotes the Gâteaux derivative with respect to the mass of households at point s, which itself is a function of the entire distribution Φ.[29] For notational simplicity we will often omit the dependence on Φ and the time subscripts that denote the dependence on the path of rates.

We will use these tools to solve the planner's problem (27) using Lagrangian methods. Let $\lambda(s)$ denote the multiplier on the no-equity issuance constraint (44), $\mu(s)$ denote the multiplier on the collateral constraint (45), and Λ denote the multiplier on the non-rivalry

[29] A more explicit analogy with partial derivatives may be useful. Suppose that the state space s lay on a finite grid with N points. Then the distribution $\Phi(s)$ would be an $N\times 1$ vector, and the value function $W(\Phi):\mathbb{R}^N \to 1$. In this case, the partial derivative $\frac{\partial W}{\partial\Phi(s_i)}:\mathbb{R}^N\to 1$ is a function of Φ as well.

externality (46). We will directly plug in the definitions of consumption (43) and the transition function for the distribution (32). With all this notation in hand, the Lagrangian is

$$\mathcal{L} = \frac{C^{1-\sigma}-1}{1-\sigma} + \int \lambda(s)\left((Az)^{1-\alpha}k^{\alpha} + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1+r_t}\right)ds$$
$$+ \int \mu(s)\left(\theta k'(s) - b'(s)\right)ds + \Lambda\left[\left(\int z\Phi(s)ds\right)^{a} - A\right] + \beta W(T(\Phi; k'(\cdot), i(\cdot), b'(\cdot))),$$

where, again, it is understood that C and $T(\Phi; k'(\cdot), i(\cdot), b'(\cdot))$ stand in for (43) and (32).

We proceed in two steps. First, subsection E.1 takes the first-order conditions with respect to all the planner's choices. Second, subsection E.2 characterizes those choices in terms of the marginal social value function from Proposition 4 in the main text.

E.1 First Order Conditions

We analyze each first-order condition separately.

Aggregate productivity The FOC with respect to aggregate productivity is

$$C^{-\sigma}\left[\int (1-\alpha)A^{-\alpha}z^{1-\alpha}k^{\alpha}\Phi(s)ds - (1-\pi_d)\int zi(s)\Phi(s)ds\right]$$
$$+ \int \lambda(s)\left[(1-\alpha)A^{-\alpha}z^{1-\alpha}k^{\alpha} - zi(s)\right]ds = \Lambda. \tag{33}$$

Going forward, it will be convenient to work with the transformed multipliers $\widetilde{\lambda}(s) = \frac{\lambda(s)}{\Phi(s)(1-\pi_d)C^{-\sigma}}$ and $\widetilde{\Lambda} = \frac{\Lambda}{C^{-\sigma}}$.[30] Plugging these in and simplifying yields

$$\widetilde{\Lambda} = \pi_d \underbrace{\int (1-\alpha)A^{-\alpha}z^{1-\alpha}k^{\alpha}\Phi(s)ds}_{=\widetilde{\Lambda}^{\text{exit}}} + (1-\pi_d)\underbrace{\int (1+\widetilde{\lambda}(s))\left[(1-\alpha)A^{-\alpha}z^{1-\alpha}k^{\alpha} - zi(s)\right]\Phi(s)ds}_{=\widetilde{\Lambda}^{\text{cont}}}. \tag{34}$$

Innovation The FOC with respect to innovation at a particular point $i(s)$ is

$$C^{-\sigma}(1-\pi_d)Az\Phi(s) + \lambda(s)Az = \beta\int \frac{\delta W(\Phi')}{\delta\Phi'(s')}\frac{\delta T(s')}{\delta i(s)}ds'.$$

[30] Of course, this transformed multiplier $\widetilde{\lambda}(s)$ is only defined for points with a positive mass of firms.

The LHS is the planner's marginal cost of higher innovation $i(s)$, which reduces consumption and tightens the no-equity issuance constraint for firm-type s. The RHS is the marginal benefit, which captures how higher innovation affects the distribution of productivity in the next period. To keep the notation manageable, we denote $T(s') = T(\Phi; k'(\cdot), i(\cdot), b'(\cdot))(s') = \Phi'(s')$. The integral is the functional-derivative extension of the chain rule: a change in $i(s)$ affects the mass of firms at each point in the state space in the next period $T(s')$, and each of those marginal changes affects the social welfare function $W(\Phi')$.

We can simplify the $\frac{\delta T(s')}{\delta i(s)}$ terms using the definition of the transition function (32). In particular, marginal changes in $i(s)$ only affect the transition function through changing the probability of success, not changing the value of the state conditional on success. Therefore, we have

$$\frac{\delta T(s')}{\delta i(s)} = \left\{ \begin{matrix} (1-\pi_d)\eta'(i(s))p(\varepsilon)\Phi(s) \text{ if } s' = (ze^{\Delta}e^{\varepsilon}, k'(s), b'(s)), \\ -(1-\pi_d)\eta'(i(s))p(\varepsilon)\Phi(s) \text{ if } s' = (ze^{\varepsilon}, k'(s), b'(s)) \\ 0 \text{ otherwise} \end{matrix} \right\}$$

Plugging this into the FOC gives

$$C^{-\sigma}(1-\pi_d)Az\Phi(s) + \lambda(s)Az = \beta(1-\pi_d)\eta'(i(s))\Phi(s) \left[\begin{matrix} \int \frac{\delta W(\Phi')}{\delta\Phi'(ze^{\Delta}e^{\varepsilon}, k'(s), b'(s))} p(\varepsilon)d\varepsilon \\ - \int \frac{\delta W(\Phi')}{\delta\Phi'(ze^{\varepsilon}, k'(s), b'(s))} p(\varepsilon)d\varepsilon \end{matrix} \right]$$

Finally, dividing by $C^{-\sigma}(1-\pi_d)\Phi(s)$ and using our definition of $\widetilde{\lambda}(s)$ from above gives

$$Az(1+\widetilde{\lambda}(s)) = \frac{\beta}{C^{-\sigma}}\eta'(i(s)) \left[\int \frac{\delta W(\Phi')}{\delta\Phi'(ze^{\Delta}e^{\varepsilon}, k'(s), b'(s))} p(\varepsilon)d\varepsilon - \int \frac{\delta W(\Phi')}{\delta\Phi'(ze^{\varepsilon}, k'(s), b'(s))} p(\varepsilon)d\varepsilon \right]. \quad (35)$$

Investment The FOC for capital accumulation at a particular point $k'(s)$ is

$$C^{-\sigma}(1-\pi_d)\Phi(s) + \lambda(s) = \theta\mu(s) + \beta \int \frac{\delta W(\Phi')}{\delta\Phi'(s')} \frac{\delta T(s')}{\delta k'(s)} ds'.$$

The derivatives of next period's value functions are more complicated than for innovation because a marginal change in $k'(s)$ affects the value of the state s' in the next period.

Assuming we can swap the order of differentiation, we can write

$$\int \frac{\delta W(\Phi')}{\delta \Phi'(s')} \frac{\delta T(s')}{\delta k'(s)} ds' = \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(s')} T(s') ds'.$$

Plugging in the definition of the transition function, noting only the part of the transition function from incumbents will matter for the derivatives, and swapping the order of integration gives

$$\int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(s')} T(s') ds' = (1-\pi_d) \int\int\int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(s')} \begin{bmatrix} (\mathbb{1}\{k' = k'(s)\} \times \mathbb{1}\{b' = b'(s)\}) \times \\ (\eta(i(s)) \mathbb{1}\{z' = ze^{\Delta} e^{\varepsilon}\} + (1-\eta(i(s))) \mathbb{1}\{z' = ze^{\varepsilon}\}) \end{bmatrix} p(\varepsilon) \Phi(s) ds ds' d\varepsilon.$$

Using only the initial state s under consideration and eliminating the values of the future state variables s' with zero probability, the integral becomes

$$(1-\pi_d) \left[\eta(i(s)) \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\Delta} e^{\varepsilon}, k'(s), b'(s)))} p(\varepsilon) d\varepsilon + (1-\eta(i(s))) \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\varepsilon}, k'(s), b'(s)))} p(\varepsilon) d\varepsilon \right] \Phi(s).$$

Finally, we will plug this into the FOC, and as usual divide by $C^{-\sigma}(1-\pi_d)\Phi(s)$ to get

$$1 + \widetilde{\lambda}(s) = \theta \widetilde{\mu}(s) + \frac{\beta}{C^{-\sigma}} \begin{bmatrix} \eta(i(s)) \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\Delta} e^{\varepsilon}, k'(s), b'(s)))} p(\varepsilon) d\varepsilon + \\ (1-\eta(i(s))) \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\varepsilon}, k'(s), b'(s)))} p(\varepsilon) d\varepsilon \end{bmatrix} \tag{36}$$

where $\widetilde{\mu}(s) = \frac{\mu(s)}{C^{-\sigma}(1-\pi_d)\Phi(s)}$.

Borrowing The FOC for borrowing at a particular point $b'(s)$ is

$$\frac{\lambda(s)}{1+r_t} = \mu(s) - \beta \int \frac{\delta W(\Phi')}{\delta \Phi'} \frac{\delta T(s')}{\delta b'(s)} ds'$$

As with capital, we can write the integral term as

$$\int \frac{\delta}{\delta b'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(s')} T(s') ds' = (1-\pi_d) \int\int\int \frac{\delta}{\delta b'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(s')} \begin{bmatrix} (\mathbb{1}\{k' = k'(s)\} \times \mathbb{1}\{b' = b'(s)\}) \times \\ (\eta(i(s)) \mathbb{1}\{z' = ze^{\Delta} e^{\varepsilon}\} + (1-\eta(i(s))) \mathbb{1}\{z' = ze^{\varepsilon}\}) \end{bmatrix} p(\varepsilon) \Phi(s) ds ds' d\varepsilon.$$

And as in the case with capital, this integral becomes

$$(1-\pi_d)\left[\eta(i(s))\int \frac{\delta}{\delta b'(s)}\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\Delta}e^{\varepsilon},k'(s),b'(s)))}p(\varepsilon)d\varepsilon+(1-\eta(i(s)))\int \frac{\delta}{\delta b'(s)}\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\varepsilon},k'(s),b'(s)))}p(\varepsilon)d\varepsilon\right]\Phi(s).$$

Plugging this into the FOC and dividing by $C^{-\sigma}(1-\pi_d)\Phi(s)$ yields

$$\frac{\widetilde{\lambda}(s)}{1+r_t}=\widetilde{\mu}(s)-\frac{\beta}{C^{-\sigma}}\begin{bmatrix}\eta(i(s))\int \frac{\delta}{\delta b'(s)}\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\Delta}e^{\varepsilon},k'(s),b'(s)))}p(\varepsilon)d\varepsilon+\\(1-\eta(i(s)))\int \frac{\delta}{\delta b'(s)}\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\varepsilon},k'(s),b'(s)))}p(\varepsilon)d\varepsilon\end{bmatrix} \tag{37}$$

E.2 Marginal Social Value Functions

The optimal choices to the planner's problem are given the FOCs (34), (35), (36), and (37), together with the complementarity slackness conditions. In order to arrive at the results in Proposition 4, we now use the envelope theorem to get a recursive expression for the marginal social value function $\frac{\delta W(\Phi)}{\delta\Phi(s)}$.

Differentiating the RHS of the planner's objective at the optimal policies results in

$$\begin{aligned}\frac{\delta W(\Phi)}{\delta\Phi(s)}=&C^{-\sigma}\left[(Az)^{1-\alpha}k^{\alpha}+(1-\delta)k-(1-\pi_d)\left(k'(s)+Azi(s)\right)\right]+\Lambda a\left(\int z\Phi(s)ds\right)^{a-1}z\\&+\beta\int\frac{\delta W(\Phi')}{\delta\Phi'(s')}\frac{\delta T(s')}{\delta\Phi(s)}ds'.\end{aligned}$$

From the definition of the transition function (32), we have

$$\int\frac{\delta W(\Phi')}{\delta\Phi'(s')}\frac{\delta T(s')}{\delta\Phi(s)}ds'=(1-\pi_d)\left[\eta(i(s))\int\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\Delta}e^{\varepsilon},k'(s),b'(s))}p(\varepsilon)d\varepsilon+(1-\eta(i(s)))\int\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\varepsilon},k'(s),b'(s))}p(\varepsilon)d\varepsilon\right].$$

We now define $\omega(s;\Phi)=\frac{\delta W(\Phi)}{\delta\Phi(s)}$ to be the marginal social value function in the direction of $\Phi(s)$. Plugging this into the two equations above and slightly rearranging, we have

$$\begin{aligned}\omega(s;\Phi)=&\pi_d C^{-\sigma}\left[(Az)^{1-\alpha}k^{\alpha}+(1-\delta)k\right]+(1-\pi_d)C^{-\sigma}\left[(Az)^{1-\alpha}k^{\alpha}+(1-\delta)k-k'(s)-Azi(s)\right]\\&+\Lambda a\left(\int z\Phi(s)ds\right)^{a-1}z+\beta(1-\pi_d)\mathbb{E}^{\varepsilon}\left[\eta(i(s))\omega(s';\Phi')+(1-\eta(i(s)))\omega(s';\Phi')\right],\end{aligned}$$

where $\mathbb{E}^{\varepsilon}[\omega(s';\Phi')] = \int \omega(s';\Phi')p(\varepsilon)d\varepsilon$ takes the expectation over idiosyncratic shocks ε.

We now define $\widetilde{\omega}(s;\Phi) = \frac{\omega(s;\Phi)}{C^{-\sigma}}$. Plugging this into the equation above yields

$$\begin{aligned}\widetilde{\omega}(s;\Phi) =&\pi_d\left[(Az)^{1-\alpha} + (1-\delta)k + \widetilde{\Lambda}^{\text{exit}} a\left(\int z\Phi(s)ds\right)^{a-1} z\right] + \\ &+(1-\pi_d)\left[\begin{array}{l}(Az)^{1-\alpha}k^{\alpha} + (1-\delta)k - k'(s) - Azi(s) + \widetilde{\Lambda}^{\text{cont}} a\left(\int z\Phi(s)ds\right)^{a-1} z \\ +\beta\left(\frac{C'}{C}\right)^{-\sigma}\mathbb{E}^{\varepsilon}\left[\eta(i(s))\widetilde{\omega}(s';\Phi') + (1-\eta(i(s)))\widetilde{\omega}(s';\Phi')\right]\end{array}\right].\end{aligned} \tag{38}$$

We are finally in a position to derive the equations in Proposition 4 from the main text. Let time subscripts denote the optimal value and policy functions conditional on the optimal path of the distribution $\Phi(s)$. Then, let

$$\widehat{\omega}_t(s) = \widehat{\omega}(s;\Phi_t) - b'_{t-1}(s) + (1-\pi_d)\frac{b'_t(s)}{1+r_t} + \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}(1-\pi_d)\left(-b'_t(s) + (1-\pi_d)\frac{b'_{t+1}(s)}{1+r_{t+1}}\right) + \ldots$$

be the planner's social marginal value function plus the path of borrowing and debt repayments starting from period t. Plugging this into (38) gives the *augmented Bellman equation*

$$\begin{aligned}\widehat{\omega}_t(s) =&\pi_d\left[(Az)^{1-\alpha} + (1-\delta)k - b + \widetilde{\Lambda}^{\text{exit}} a\left(\int z\Phi(s)ds\right)^{a-1} z\right] + \\ &+(1-\pi_d)\left[\begin{array}{l}(Az)^{1-\alpha}k^{\alpha} + (1-\delta)k - b - k'(s) - Azi(s) + \widetilde{\Lambda}^{\text{cont}} a\left(\int z\Phi(s)ds\right)^{a-1} z \\ +\frac{b'(s)}{1+r_t} + \beta\left(\frac{C'}{C}\right)^{-\sigma}\mathbb{E}^{\varepsilon}\left[\eta(i(s))\widehat{\omega}_t(s') + (1-\eta(i(s)))\widehat{\omega}_t(s')\right]\end{array}\right].\end{aligned} \tag{39}$$

To keep notation even simpler, define $\widehat{\Lambda}^{\text{exit}} = \widetilde{\Lambda}^{\text{exit}} a\left(\int z\Phi(s)ds\right)^{a-1}$ and similarly $\widehat{\Lambda}^{\text{cont}} = \widetilde{\Lambda}^{\text{cont}} a\left(\int z\Phi(s)ds\right)^{a-1}$. In addition, let $\mathbb{E}_t$ denote the expectation over both the innovation shock and the idiosyncratic ε shocks, as in the main text. Finally, let $\widehat{\omega}_t^{\text{exit}}$ denote the terms inside the first set of brackets in (39) and let $\widehat{\omega}_t^{\text{cont}}$ second set of brackets in (39). Then we have $\widehat{\omega}_t(s) = \pi_d\widehat{\omega}_t(s)^{\text{exit}} + (1-\pi_d)\widehat{\omega}_t(s)^{\text{cont}}$, where

$$\widehat{\omega}_t^{\text{cont}}(s) = (Az)^{1-\alpha}k^{\alpha} + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1+r_t} + \widehat{\Lambda}^{\text{cont}} z + \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\mathbb{E}_t\left[\widehat{\omega}_{t+1}(s')\right] \tag{40}$$

This Bellman-like equation (40) is similar to the augmented Bellman equation (12) from Proposition 4 except that (40) is evaluated at the planner's optimal policies. Therefore, it remains to show that the planner's policies maximize the RHS of Bellman operator implied by the RHS of (40) subject to the constraints $d \geq 0$ and $b' \leq \theta k'$. But inspection of the FOCs we derived above shows that this is the case.

F Modeling the Corporate Tax System

We model the structure of the U.S. corporate tax code before the Tax Cuts and Jobs Act (TCJA 2017), and then consider the long-run effects of implementing the TCJA 2017. We assume firms pay a linear tax rate τ on their revenues net of tax deductions. Firms can fully deduct innovation expenditures in the period in which they occur, but investment expenditures must be gradually deducted over time according to the tax depreciation schedule.[31] Following Winberry (2021), we assume the tax deduction schedule follows a geometric depreciation process with tax depreciation rate $\widehat{\delta}$ (which may differ from economic depreciation δ). Each period, firms inherit a stock of depreciation allowances $\widehat{k}_{jt}$ from past investments and deduct the fraction $\widehat{\delta}$ of those depreciation allowances from their tax bill. In addition, firms deduct the same fraction $\widehat{\delta}$ of new investment $k_{jt+1} - (1-\delta)k_{jt}$ from their tax bill as well. Therefore, their total tax bill in a given period is

$$\tau \times \left(y_{jt} - A_t z_{jt} i_{jt} - \widehat{\delta} \left[\widehat{k}_{jt} + (k_{jt+1} - (1-\delta)k_{jt}) \right] \right).$$

The firm carries the un-deducted portion of its investments into the next period: $\widehat{k}_{jt+1} = (1-\widehat{\delta}) \left[\widehat{k}_{jt} + (k_{jt+1} - (1-\delta)k_{jt}) \right]$.

In principle, we would need two new state variables, $\widehat{k}_{jt}$ and k_{jt}, in order to forecast the evolution the stock of depreciation allowances $\widehat{k}_{jt+1}$. However, we are able to bypass these additional states using the following simplifying assumption.

***Proposition* 3.** *Suppose that firms can borrow against future tax deductions at the risk-free rate r_t. Then the tax depreciation schedule only affects firm decisions through the present*

[31] Empirically, R&D expenditures are typically fully deducted because they primarily reflect labor costs.

value of tax deductions per unit of investment:

$$\widehat{\zeta}_t = \sum_{s=0}^{\infty} \left(\prod_{p=0}^{s} \frac{1}{1 + r_{t+p}} \right) (1 - \widehat{\delta})^s. \tag{41}$$

This present value alters the effective after-tax price of capital:

$$v_t^{cont}(z, n) = \max_{k', i, b'} n - (1 - \tau\widehat{\zeta}_t)k' - (1 - \tau)A_t zi + \frac{b'}{1 + r_t} + \frac{1}{1 + r_t}\mathbb{E}_t \left[v_{t+1}(z', n') \right] \text{ s.t. } d \geq 0 \text{ and } b' \leq \theta k',$$

where $n' = (1 - \tau)(A_t z')^{1-\alpha}(k')^{\alpha} + (1 - \tau\widehat{\zeta}_t)(1 - \delta)k - b'$.

Proof. The key insight of our proof is that borrowing against the stream of future tax deductions is equivalent to selling a claim on this stream to households. Since the claim is risk-free, the household is willing to pay its present value $\tau\widehat{\zeta}_t \times (k_{jt+1} - (1 - \delta)k_{jt})$. Hence, each unit of investment produces $\tau\widehat{\zeta}_t$ of additional resources to the firm, lowering its after-tax price by that amount.

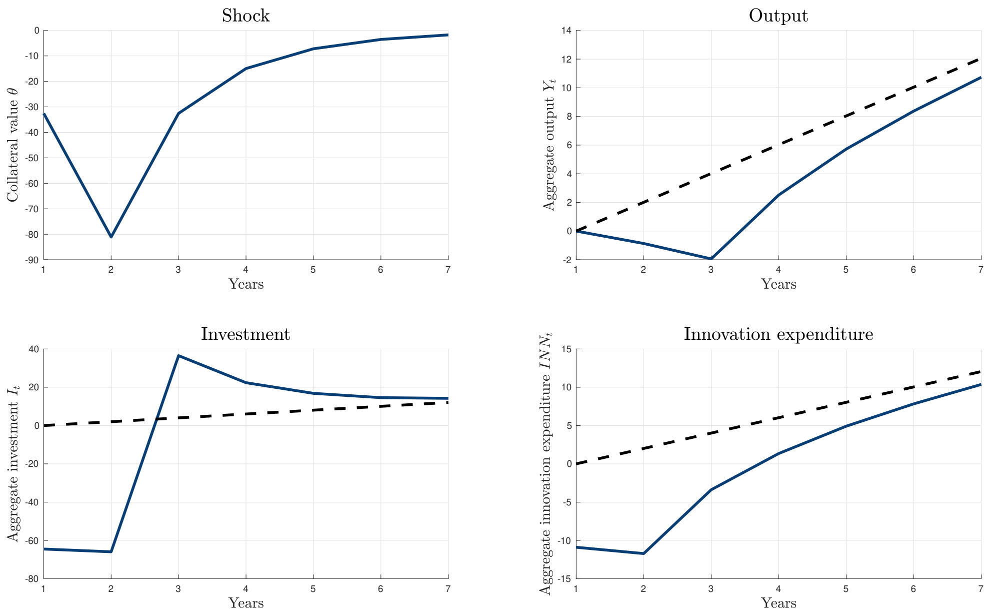
The financially constrained firms from Proposition 1 (with a positive financial wedge $\lambda_t(z, n) > 0$) will strictly prefer to sell the claim because their shadow value of funds is higher than the household's value of funds. However, financially unconstrained firms (with no financial wedge $\lambda_t(z, n) = 0$) will be indifferent between selling the claim or not because they value funds the same as the household. However, one can show that in this case, the present value of the tax deductions affects firms decisions because they are indifferent over the timing (technically, their value function is linearly separable in the tax deductions; see Winberry (2021)). ■

This proposition allows us to model both temporary investment tax incentives and permanent tax reforms using changes in the present value $\widehat{\zeta}_t$.

G Additional Results

This section contains two additional quantitative results mentioned in the main text. First, Section G.1 shows that temporary financial shocks have relatively transitory effects in our

FIGURE 19: Aggregate Transition Paths, Financial Shock (Partial Equilibrium)



Notes: aggregate transition paths following an unexpected tightening of the collateral constraint θ_t. Top left panel plots the path of θ_t. Remaining panels plot aggregate output, investment, and innovation expenditures in log-deviations from the initial period. Dashed black lines are the growth trajectory in the initial market BGP. Real interest rate $r_t = r^*$ is kept fixed at its value in the initial market BGP.

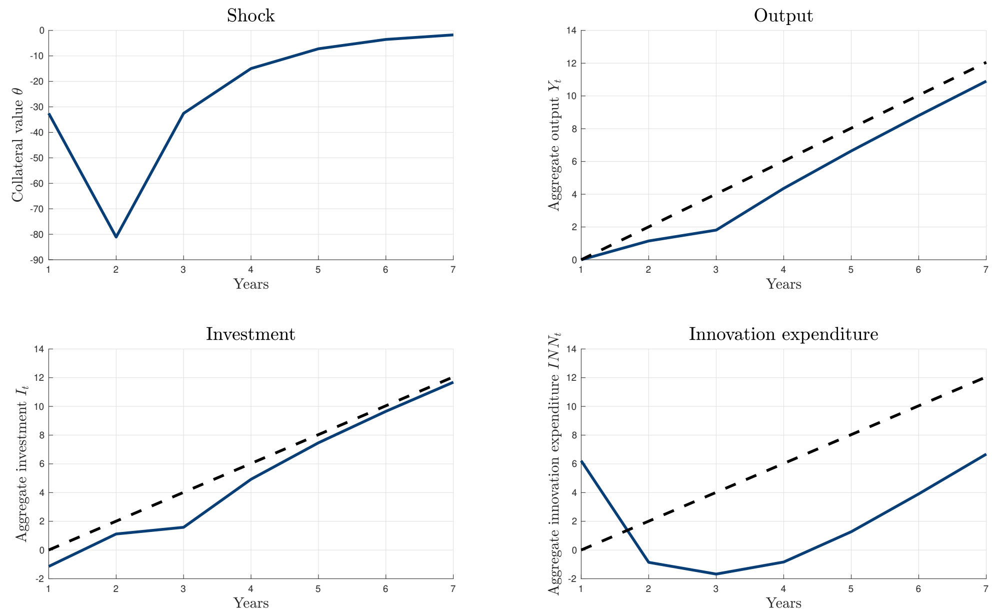
model. Second, Section G.2 studies the effect of full expensing in the partial equilibrium version of the model with fixed interest rates $r_t = r^*$.

G.1 Small Growth Effects of Financial Shocks

Figure 19 plots the effects of a transitory tightening of the financial constraint θ_t in partial equilibrium, i.e. holding the real interest rate $r_t = r^*$ fixed at its initial value. The path of the shock is in the top plotted in the top left graph. The shock reduces available financing, which directly lowers investment and, to some extent, innovation. However, once the shock reverts back to the steady state, investment and innovation recover relatively quickly.

Figure 20 performs the same exercise in general equilibrium, i.e. when the real interest rate r_t is endogenously determined. The real interest rate r_t declines given the fall in investment demand, which dampens the fall in investment. More strikingly, by raising the relative return to innovation, the lower real interest rate actually *raises* the innovation rate. Hence,

FIGURE 20: Aggregate Transition Paths, Financial Shock (General Equilibrium)



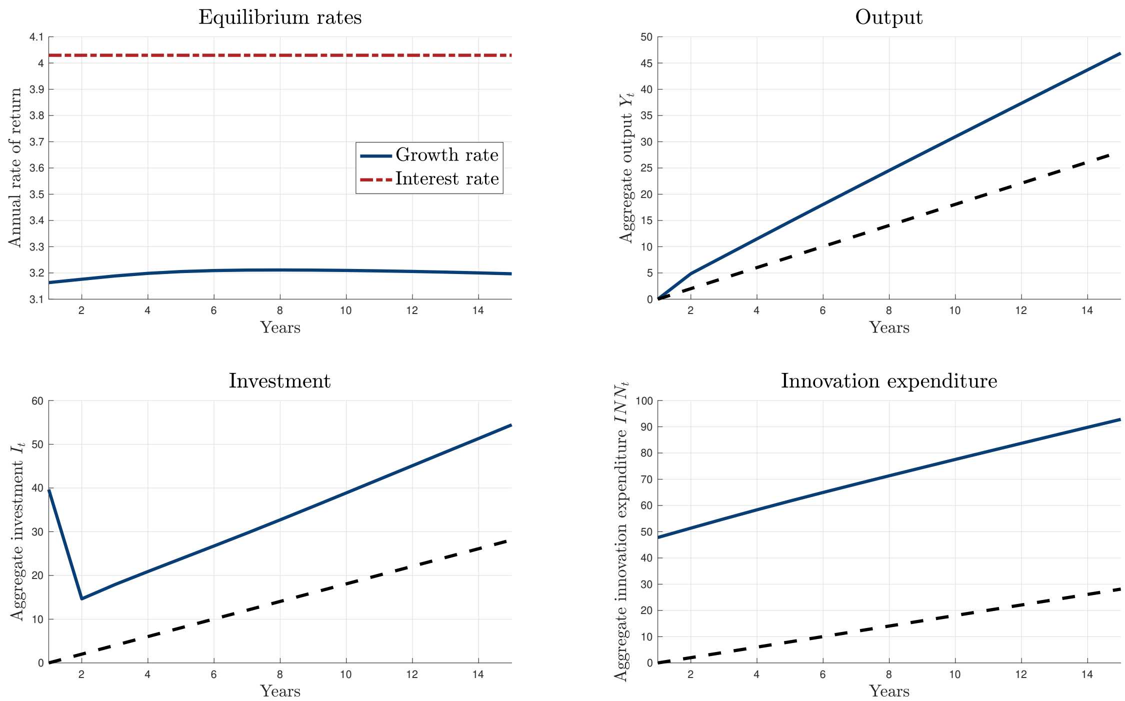
Notes: aggregate transition paths following an unexpected tightening of the collateral constraint θ_t. Top left panel plots the path of θ_t. Remaining panels plot aggregate output, investment, and innovation expenditures in log-deviations from initial period. Dashed black lines are the growth trajectory in the initial market BGP.

the effects of the financial shock are even more short-lived than in partial equilibrium.

G.2 Full Expensing in Partial Equilibrium

Figure 21 plots the effects of full expensing in partial equilibrium, i.e. holding the real interest rate fixed at its initial level $r_t = r^*$. We note two quantitative differences from our general equilibrium results in the main text. First, as described in the main text, the response of innovation is always positive without the dampening effect of higher real interest rates. Second, the overall response of both investment and innovation are much higher in partial equilibrium than in general equilibrium. This occurs because, without adjustment costs, our financially unconstrained firms are extremely sensitive to price movements.

FIGURE 21: The Effects of Full Investment Expensing in Partial Equilibrium



Notes: transition path following an unexpected, permanent decline in the relative price of capital $1 - \zeta_t$ of the size equivalent to full expensing of investment, starting from the initial period. Dashed lines correspond to the paths of investment, output, and innovation in the initial market BGP. Solid lines correspond to their actual paths in response to the change in the relative price of capital. Real interest rate $r_t = r^*$ is kept fixed at its value in the initial market BGP.

H Extended Model with Labor

In this appendix, we add labor to the model and study pecuniary externalities arising from the real wage. We also extend the planner's problem from Appendix E to incorporate pecuniary externalities from the real interest rate as well.

Adding labor extends the model in two ways. First, as discussed in the main text, the production function becomes $y_{jt} = (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha} \ell_{jt}^{\nu}$, where ℓ_{jt} is the labor used in production by firm j and $\alpha + \nu < 1$. Second, we incorporate labor supply into the household's preferences by assuming that the utility function is

$$\sum_{t=0}^{\infty} \beta^t \left[\log C_t - \chi \frac{L_t^{1+\psi}}{1+\psi} \right],$$

where χ is a scale parameter and ψ^{-1} is the Frisch elasticity of labor supply.[32]

We now describe how this extended model affects our positive and normative results.

H.1 Positive Results

Adding labor does not significantly alter our positive results; it simply leads to a re-interpretation of the production function in the main text. To see this, note that firms' optimal labor demand is purely static and is therefore independent of their net worth:

$$\max_{\ell_{jt}} (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha} \ell_{jt}^{\nu} - w_t \ell_{jt} \quad \implies \ell_{jt} = \left(\frac{\nu (A_t z_{jt})^{1-\alpha} k_{jt}^{\alpha}}{w_t} \right)^{\frac{1}{1-\nu}}$$

Now define variable profits $\pi_{jt} = y_{jt} - w_t \ell_{jt}$. Plugging in the above expression for optimal labor demand and simplifying yields

$$\pi_{jt} = \widetilde{\nu} \, (A_t z_{jt})^{\frac{1-\alpha}{1-\nu}} \, w_t^{-\frac{\nu}{1-\nu}} k_{jt}^{\frac{\alpha}{1-\nu}}.$$

where $\widetilde{\alpha} = \frac{\alpha}{1-\nu}$ and $\widetilde{\nu} = \nu^{\frac{\nu}{1-\nu}} - \nu^{\frac{1}{1-\nu}}$.

The firm's problem in this extended model is isomorphic to our previous model using the new definition of net worth: $n_{jt} = \pi_{jt} + (1-\delta) k_{jt} - b_{jt}$. Importantly, net worth still grows with Z_t, facilitating the same detrending as in our baseline model. Specifically, it is easy to guess and verify that the real wage w_t scales with Z_t, which implies that the first two terms grow with $Z_t^{\frac{1-\alpha-\nu}{1-\nu}} = Z_t^{1-\frac{\alpha}{1-\nu}}$. But since capital grows with Z_t, the term involving capital grows with $Z_t^{\frac{\alpha}{1-\nu}}$. Putting these two observations together, variable profits grows with $Z_t^{1-\frac{\alpha}{1-\nu}} Z_t^{\frac{\alpha}{1-\nu}} = Z_t$.

The equilibrium of this extended model is the same as in our baseline model, except that we add the real wage w_t as another equilibrium price and add the labor market as another market clearing condition:

$$\left(\frac{w_t C_t^{-1}}{\chi} \right)^{\frac{1}{\psi}} = \int \ell_{jt} dj.$$

[32] Given these additively separable preferences over consumption and labor supply, balanced growth requires log utility over consumption.

H.2 Normative Results

We set up the planner's problem analogously to that in Appendix E, but extend that treatment to include pecuniary externalities from both the real wage and the real interest rate. Adding the real interest rate requires that the planner includes the household's Euler equation as an additional constraint. We incorporate this constraint using the promised utility approach. In this case, the planner's state variable for any period $t \geq 0$ is the distribution of firms, $\Phi(s)$, and the promised marginal utility of consumption, M (inherited from past consumption and interest rate choices). In the initial period $t = 0$, the planner does not inherit a promised value of marginal utility M.

Planner's Problem The planner chooses $\ell^{\text{exit}}(s), \ell^{\text{cont}}(s), k'(\cdot), b'(\cdot), i(\cdot), C, r, w, A$ in order to maximize

$$W(\Phi, M) = \log C - \chi \frac{L^{1+\psi}}{1+\psi} + \beta W(T(\Phi; k'(\cdot), b'(\cdot), i(\cdot)), \frac{M}{1+r}). \tag{42}$$

Note that the planner may choose different levels of employment for exiting firms $\ell^{\text{exit}}(s)$ and continuing firms $\ell^{\text{cont}}(s)$. The transition function for the distribution of firms $T(\Phi; k'(\cdot), b'(\cdot), i(\cdot))$ is the same as in Appendix E.

The planner faces the following constraints (with associated Lagrange multipliers)

$$C = \pi_d \int \left[(Az)^{1-\alpha} k^\alpha \ell^{\text{exit}}(s)^\nu + (1-\delta)k\right] \Phi(s)ds + (1-\pi_d) \int \left[(Az)^{1-\alpha} k^\alpha \ell^{\text{cont}}(s)^\nu + (1-\delta)k\right] \Phi(s)ds$$
$$- (1-\pi_d) \int [k'(s) + Azi(s)] \Phi(s)ds - \pi_d k_0 \qquad (\times\kappa) \tag{43}$$

$$(Az)^{1-\alpha} k^\alpha \ell^{\text{cont}}(s)^\nu - w\ell^{\text{cont}}(s) + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1+r} \geq 0 \forall s \; (\times \mu(s)(1-\pi_d)\Phi(s)) \tag{44}$$

$$b'(s) \leq \theta k'(s) \forall s \qquad (\times \lambda(s)(1-\pi_d)\Phi(s)) \tag{45}$$

$$A = \left(\int z\Phi(s)dz\right)^a \qquad (\times \Lambda_A) \tag{46}$$

$$M = C^{-\sigma} \qquad (\times \Lambda_C) \tag{47}$$

$$wC^{-1} = \chi L^\psi \qquad (\times \Lambda_L) \tag{48}$$

It is understood that we will substitute $L = \pi_d \int \ell^{\text{exit}}(s)\Phi(s)ds + (1-\pi_d) \int \ell^{\text{cont}}(s)\Phi(s)ds$ where necessary.

***Proposition* 4.** *In the constrained-efficient equilibrium, incumbent firms' decisions solve the augmented Bellman equation*

$$\omega_t^{cont}(s) = \max_{\ell,k',b',i} (Az)^{1-\alpha}k^\alpha\ell^\nu - \widetilde{w}_t\ell + (1-\delta)k - b - k' - Azi + \frac{b'}{1+r_t}$$
$$+ \widetilde{\Lambda}_{At}z + \frac{1}{1+\widetilde{r}_t}\mathbb{E}_t[\omega_{t+1}(s')] \text{ s.t. } b' \leq \theta k' \text{ and}$$
$$d = (Az)^{1-\alpha}k^\alpha\ell^\nu - w_t\ell + (1-\delta)k - b - k' - Azi + \frac{b'}{1+r_t} \geq 0.$$

The strength of the non-rivalry externality is analogous to (13) *in the main text:*

$$\widetilde{\Lambda}_{At} = a\left(\int z\Phi_t(s)ds\right)^{a-1} \times \begin{bmatrix} \pi_d \int (1-\alpha)A_t^{-\alpha}z^{1-\alpha}k^\alpha\ell_t^{exit}(s)^\nu\Phi_t(s)ds \\ + (1-\pi_d)\int(1+\widetilde{\lambda}_t(s))\left[(1-\alpha)A_t^{-\alpha}z^{1-\alpha}k^\alpha n_t^{cont}(s)^\nu - zi_t(s)\right]\Phi_t(s)ds \end{bmatrix}$$

where $\widetilde{\lambda}_t(s)$ is the firm's shadow value of funds. The pecuniary externalities are captured in the social prices $\widetilde{r}_t$ and $\widetilde{w}_t$, which are related to the market prices r_t and w_t through

$$\widetilde{w}_t = w_t\left(1+\widetilde{\Lambda}_{Lt}\psi L_t^{-1}\right)$$
$$\frac{1}{1+\widetilde{r}_t} = \frac{1}{1+r_t}\frac{1+C_{t+1}^{-1}\widetilde{\Lambda}_{Ct+1} - \widetilde{\Lambda}_{Lt+1}w_{t+1}C_{t+1}^{-1}}{1+C_t^{-1}\widetilde{\Lambda}_{Ct} - \widetilde{\Lambda}_{Lt}w_tC_t^{-1}}$$

where $\widetilde{\Lambda}_{Ct}$ and $\widetilde{\Lambda}_{Lt}$ measure the strength of the pecuniary externalities through

$$\Omega_{t+1} = -\frac{1-\pi_d}{1+\widetilde{\Lambda}_{Ct}C_t^{-1} - \widetilde{\Lambda}_{Lt}w_tC_t^{-1}}\int\widetilde{\lambda}_t(s)b_t'(s)\Phi_t(s)ds, \text{ where } \Omega_{t+1} \text{ is from}$$
$$\Omega_t = \Lambda_{Ct} + \frac{1}{1+r_t}\Omega_{t+1} \text{ starting from } \Omega_0 = 0$$
$$\widetilde{\Lambda}_{Lt} = \frac{1-\pi_d}{1+\widetilde{\Lambda}_{Ct}C_t^{-1} - \widetilde{\Lambda}_{Lt}w_tC_t^{-1}}\int\widetilde{\lambda}_t(s)\ell_t^{cont}(s)\Phi_t(s)ds.$$

Proof We now take the first order conditions of the planner's problem with respect to their choice variables. The conditions for aggregate productivity, capital accumulation, innovation, and borrowing are the same as in Appendix E.

The first order condition for the real interest rate r is

$$\frac{\partial W(\Phi', M')}{\partial M'} = -\frac{(1-\pi_d)\kappa}{M}\int\widetilde{\lambda}(s)b'(s)\Phi(s)ds \tag{49}$$

where $\widetilde{\lambda}(s) = \lambda(s)/\kappa$. A higher promised marginal utility lowers the planner's value because it raises the real interest rate, which has a negative pecuniary externality to the extent that firms are financially constrained.

The first order condition for consumption differs for $t = 0$ and $t \geq 1$ because the planner does not inherit a costate M in $t = 0$. For $t \geq 1$, the first order condition is

$$\kappa = C^{-1}\left(1 + C^{-1}\Lambda_C - \Lambda_N w C^{-1}\right). \tag{50}$$

We will show below that $\Lambda_C \leq 0$ and $\Lambda_N \geq 0$, i.e. the planner discount dates in which the pecuniary externalities are more binding.

For $t = 0$, the planner does not face the constraint that $M = C^{-1}$, but understands that its choice of C_0 enters the continuation value state variable through $C_1^{-1} = \frac{C_0^{-1}}{\beta(1+r)}$. The first order condition is

$$\kappa_0 = C_0^{-1}\left(1 - \frac{1}{1+r}\frac{\partial W(\phi_1, M_1)}{\partial M_1} - \Lambda_N w C^{-1}\right). \tag{51}$$

We will show below that these two first order conditions can be collapsed to a single condition with an appropriate initial condition for Λ_C.

The first order condition for the wage is

$$\Lambda_L C^{-1} = \kappa(1 - \pi_d)\int \widetilde{\lambda}(s)\ell^{\text{cont}}(s)\Phi(s)ds, \tag{52}$$

where $\widetilde{\Lambda}_L = \Lambda_L/\kappa$. The multiplier $\widetilde{\Lambda}_L$ measures the strength of the pecuniary externality from the wage. It is zero in the model without financial frictions, i.e. if $\widetilde{\lambda}(s) = 0$ for all s.

As described above, exiting firms and continuing firms may have different levels of labor. The first order condition for exiting firms is

$$(Az)^{1-\alpha}k^{\alpha}\nu\ell^{\text{exit}}(s)^{\nu-1} = \frac{\chi L^{\psi}}{\kappa}\left(1 + \Lambda_L\psi N^{-1}\right) \equiv \widetilde{w} \tag{53}$$

where we use the notation $\widetilde{w}$ to denote the "shadow wage" to which the planner equates these firms' marginal products. Plugging in the expression for κ from above gives

$$\widetilde{w}=\frac{\chi L^{\psi}}{C^{-1}}\frac{1+\Lambda_L\psi K^{-1}}{1+C^{-1}\Lambda_C-\Lambda_L w C^{-1}}=w\times\frac{1+\Lambda_L\psi L^{-1}}{1+C^{-1}\Lambda_C-\Lambda_L w C^{-1}} \tag{54}$$

where the second equality plugs in the labor supply FOC. Note that, if there are no financial frictions, the planner does not value the pecuniary externalities $\Lambda_L=\Lambda_C=0$ so $w=\widetilde{w}$, i.e. the shadow wage is equal to the marginal rate of substitution between consumption and leisure. On the other hand, if there are financial frictions so that higher labor imposes pecuniary externalities on constrained firms, then $\Lambda_L>0$ and (as argued below) $\Lambda_C<0$, which then implies that $\widetilde{w}>w$. The gap between the market wage and the shadow wage reflects exactly the strength of the pecuniary externalities.

The first order condition for the labor of continuing firms will also feature this shadow wage. In particular, the FOC can be simplified to

$$\widetilde{w}+\widetilde{\lambda}(s)w=(1+\widetilde{\lambda}(s))(Az)^{1-\alpha}k^{\alpha}\nu\ell^{\text{cont}}(s)^{\nu-1}. \tag{55}$$

Note that this equation is the same as for exiting firms if the firm is financially unconstrained, i.e. $\widetilde{\lambda}(s)=0$. Otherwise, the presence of financial constraints has two effects: it raises the marginal cost of hiring by tightening the no-equity issuance constraint and raises the marginal benefit of hiring by reducing the no-equity issuance constraint. In the market equilibrium, these forces would cancel so that the firm simply equates the market wage to the marginal product of labor. But the planner's shadow wage differs from the market wage which governs the tightening of financial constraints. This is why the planner chooses different levels of labor for exiting and continuing firms.

We now use the envelope theorem to get expressions for the marginal social value functions. For $t\geq 1$, the derivative of the value function with respect to marginal utility M evaluated at the optimum is

$$\frac{\partial W(\Phi,M)}{\partial M}=\Lambda_C+\frac{1}{1+r}\frac{\partial W(\Phi',M')}{\partial M'}, \tag{56}$$

i.e. the marginal social value of an additional promise of marginal utility is equated to the present value of the pecuniary externalities as summarized by Λ_C. Let $\Omega_t = \frac{\partial W(\Phi_t, M_t)}{\partial M_t}$ denote the partial derivative evaluated along the optimal path. Then the equation above becomes

$$\Omega_t = \Lambda_{Ct} + \frac{1}{1+r_t}\Omega_{t+1} \text{ for } t \geq 1. \tag{57}$$

We will now combine this with the FOCs for aggregate consumption for $t \geq 1$, (50), and the FOC for $t = 0$, (51):

$$\kappa_t = C_t^{-1}\left(1 + C_t^{-1}\Lambda_{Ct} - \Lambda_{Nt} w_t C_t^{-1}\right) \text{ for } t \geq 1$$
$$\kappa_0 = C_0^{-1}\left(1 - \frac{1}{1+r_0}\Omega_1 - \Lambda_N w_t C^{-1}\right) \text{ for } t = 0$$

Note that if we use (57) evaluated at $t = 0$ with the initial condition $\Omega_0 = 0$, then these equations can all be combined into

$$\Omega_t = \Lambda_{Ct} + \frac{1}{1+r_t}\Omega_{t+1} \text{ for } t \geq 0 \text{ with } \Omega_0 = 0 \tag{58}$$
$$\kappa_t = C_t^{-1}\left(1 + C_t^{-1}\Lambda_{Ct} - \Lambda_{Nt} w_t C_t^{-1}\right) \text{ for } t \geq 0 \tag{59}$$

The derivative of the value function with respect to $\Phi(s)$ is

$$\begin{aligned}
\frac{\delta W(\Phi, M)}{\delta \phi(s)} &= -\chi L^{\psi}\left(\pi_d \ell^{\text{exit}}(s) + (1-\pi_d)\ell^{\text{cont}}(s)\right) + \Lambda_A a\left(\int z\phi(s)ds\right)^{a-1} z + (1-\pi_d)\mu(s)\left(\theta k'(s) - b'(s)\right) \\
&+ (1-\pi_d)\lambda(s)\left((Az)^{1-\alpha}k^{\alpha}\ell^{\text{cont}}(s)^{\nu} - w\ell^{\text{cont}}(s) + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1++r}\right) \\
&+ \kappa\left(\pi_d\left[(Az)^{1-\alpha}k^{\alpha}\ell^{\text{exit}}(s)^{\nu} + (1-\delta)k\right] + (1-\pi_d)\left[(Az)^{1-\alpha}k^{\alpha}\ell^{\text{cont}}(s)^{\nu} + (1-\delta)k - k'(s) - Azi(s)\right]\right) \\
&- \Lambda_L \chi \psi L^{\psi-1} + \beta\int \frac{\delta W(\Phi', M')}{\delta \Phi'(s)}\frac{\delta T(s')}{\delta \Phi(s)}p(\varepsilon)d\varepsilon ds.
\end{aligned}$$

Rearranging gives

$$\frac{\delta W(\Phi,M)}{\delta\Phi(s)} = \pi_d \times \begin{Bmatrix} \kappa((Az)^{1-\alpha}k^{\alpha}\ell^{\text{exit}}(s)^{\nu} + (1-\delta)k) - \chi L^{\psi}(1+\psi L^{-1}\Lambda_L)\ell^{\text{exit}}(s) \\ +\Lambda_A a\left(\int z\phi(s)ds\right)^{a-1} z \end{Bmatrix} +$$

$$(1-\pi_d)\times\begin{Bmatrix} \kappa((Az)^{1-\alpha}k^{\alpha}\ell^{\text{exit}}(s)^{\nu} + (1-\delta)k - k'(s) - Azi(s)) - \chi L^{\psi}(1+\psi L^{-1}\Lambda_L)\ell^{\text{exit}}(s) + \\ \lambda(s)\left((Az)^{1-\alpha}k^{\alpha}\ell^{\text{cont}}(s)^{\nu} - w\ell^{\text{cont}}(s) + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1+r}\right) \\ \mu(s)(\theta k'(s) - b'(s)) + \Lambda_A a\left(\int z\Phi(s)ds\right)^{a-1} z \\ +\beta\left(\eta(i(s))\mathbb{E}^{\varepsilon}[\frac{\delta W(\phi',M')}{\delta\Phi'(z+\Delta+\varepsilon,k'(s),b'(s))}] + (1-\eta(i(s)))\mathbb{E}^{\varepsilon}[\frac{\delta W(\Phi',M')}{\delta\Phi'(z+\varepsilon,k'(s),b'(s))}]\right) \end{Bmatrix}.$$

Now define $\omega_t(s) = \frac{\delta W(\Phi_t,M_t)}{\delta\phi_t(s)}/\kappa_t$ to be the marginal social value relative to the multiplier κ_t, evaluated along the optimal path. Plugging this definition into the above yields

$$\omega_t(s) = \pi_d \times \left\{((Az)^{1-\alpha}k^{\alpha}\ell^{\text{exit}}(s)^{\nu} + (1-\delta)k - \widetilde{w}_t\ell^{\text{exit}}(s) + \widetilde{\Lambda}_A a\left(\int z\phi(s)ds\right)^{a-1} z\right\} +$$

$$(1-\pi_d)\times\begin{Bmatrix} (Az)^{1-\alpha}k^{\alpha}\ell^{\text{exit}}(s)^{\nu} + (1-\delta)k - k'(s) - Azi(s) - \widetilde{w}_t\ell^{\text{exit}}(s) + \\ \widetilde{\lambda}(s)\left((Az)^{1-\alpha}k^{\alpha}\ell^{\text{cont}}(s)^{\nu} - w_t\ell^{\text{cont}}(s) + (1-\delta)k - b - k'(s) - Azi(s) + \frac{b'(s)}{1+r_t}\right) + \\ \widetilde{\mu}(s)(\theta k'(s) - b'(s)) + \widetilde{\Lambda}_{At} a\left(\int z\phi(s)ds\right)^{a-1} z + \\ \beta\frac{\kappa_{t+1}}{\kappa_t}\left(\eta(i(s))\mathbb{E}^{\varepsilon}[\omega_{t+1}(z+\Delta+\varepsilon,k'(s),b'(s))] + (1-\eta(i(s)))\mathbb{E}^{\varepsilon}[\omega_{t+1}(z+\varepsilon,k'(s),b'(s))]\right) \end{Bmatrix}.$$

We are almost ready to write this in terms of a convenient dynamic programming problem. The last thing we need to do is write the social discount factor $\beta\frac{\kappa_{t+1}}{\kappa_t}$ in a more economically interpretable form. Plugging in the expression for κ_t from (59) gives

$$\beta\frac{\kappa_{t+1}}{\kappa_t} = \beta\frac{C_{t+1}^{-1}}{C_t^{-1}}\frac{1 + C_{t+1}^{-1}\Lambda_{Ct+1} - \Lambda_{Lt+1}w_{t+1}C_{t+1}^{-1}}{1 + C_t^{-1}\Lambda_{Ct} - \Lambda_{Lt}w_tC_t^{-1}}$$

Note that $\beta\frac{C_{t+1}^{-1}}{C_t^{-1}} = \frac{1}{1+r_t}$, the private discount factor. Hence, we write the social discount factor as

$$\frac{1}{1+\widetilde{r}_t} = \frac{1}{1+r_t}\frac{1 + C_{t+1}^{-1}\Lambda_{Ct+1} - \Lambda_{Lt+1}w_{t+1}C_{t+1}^{-1}}{1 + C_t^{-1}\Lambda_{Ct} - \Lambda_{Lt}w_tC_t^{-1}}. \tag{60}$$

The wedge between the private and social discount factors is again related to the pecuniary externalities. Without financial frictions, and therefore when those two externalities do not matter, then the wedge disappears and the private discount factor equals the social discount factor. If the pecuniary externalities bind more in one period or the next, the social discount factor changes.

Summing Up The planner's allocation can be characterized as the solution to the following system of equations. Incumbent firms' decisions are the solution to the augmented Bellman equation

$$\omega_t^{\text{cont}}(s) = \max_{\ell,k',b',i} (Az)^{1-\alpha}k^{\alpha}\ell^{\nu} - \widetilde{w}_t\ell + (1-\delta)k - b - k' - Azi + \frac{b'}{1+r_t} + \widetilde{\Lambda}_{At}a\left(\int z\phi(s)ds\right)^{a-1}z + \frac{1}{1+\widetilde{r}_t}\mathbb{E}_t[\omega_{t+1}(s')] \text{ s.t. } d \geq 0, b' \leq \theta k' \quad (61)$$

where $d = (Az)^{1-\alpha}k^{\alpha}\ell^{\nu} - w_t\ell + (1-\delta)k - b - k' - Azi + \frac{b'}{1+r_t}$. (Note that dividends are computed using the market wage and market interest rate, not the social ones.)

The continuation value is given by $\omega_t(s) = \pi_d\omega_t^{\text{exit}}(s) + (1-\pi_d)\omega_t^{\text{cont}}(s)$, where

$$\omega_t^{\text{exit}}(s) = \max_{\ell}(Az)^{1-\alpha}k^{\alpha}\ell^{\nu} + (1-\delta)k - \widetilde{w}_t\ell + \widetilde{\Lambda}_{At}a\left(\int z\phi(s)ds\right)^{a-1}z. \quad (62)$$

In order to compute these Bellman equations, we need to know the path of six aggregate variables: $\{g_t, w_t, C_t, \widetilde{\Lambda}_{At}, \widetilde{\Lambda}_{Ct}, \widetilde{\Lambda}_{Lt}\}$. From these variables we can compute the social prices

$$\widetilde{w}_t = w_t\left(1 + \widetilde{\Lambda}_{Nt}\psi N_t^{-1}\right) \quad (63)$$

$$\frac{1}{1+\widetilde{r}_t} = \frac{1}{1+r_t}\frac{1 + C_{t+1}^{-1}\widetilde{\Lambda}_{Ct+1} - \widetilde{\Lambda}_{Nt+1}w_{t+1}C_{t+1}^{-1}}{1 + C_t^{-1}\widetilde{\Lambda}_{Ct} - \widetilde{\Lambda}_{Nt}w_tC_t^{-1}} \quad (64)$$

where we implicitly compute L_t from $w_tC_t^{-1} = \chi L_t^{\psi}$ and r_t from $C_t^{-1} = \beta(1+r_t)C_{t+1}^{-1}$.

These six variables $\{g_t, w_t, C_t, \widetilde{\Lambda}_{At}, \widetilde{\Lambda}_{Ct}, \widetilde{\Lambda}_{Lt}\}$ must satisfy the consistency conditions

$$1 + g_t = \left(e^{\sigma_\varepsilon^2/2} \left[1 + (e^{\Delta} - 1) \frac{\int \eta(i_t(s)) z \Phi_t(s) ds}{\int z \Phi_t(s) ds} \right] \right)^{1+a} \tag{65}$$

$$\left(\frac{w_t C_t^{-1}}{\chi} \right)^{\frac{1}{\psi}} = \pi_d \int \ell_t^{\text{exit}}(s) \Phi_t(s) ds + (1 - \pi_d) \int \ell_t^{\text{cont}}(s) \Phi_t(s) ds \tag{66}$$

$$C_t = \begin{bmatrix} \pi_d \int \left[(A_t z)^{1-\alpha} k^\alpha \ell_t^{\text{exit}}(s)^\nu + (1-\delta) k \right] \Phi_t(s) ds + \\ (1 - \pi_d) \int \left[(A_t z)^{1-\alpha} k^\alpha \ell_t^{\text{cont}}(s)^\nu + (1-\delta) k - k_t'(s) - A z i_t(s) \right] \Phi_t(s) ds \end{bmatrix} \tag{67}$$

$$\widetilde{\Lambda}_{At} = \begin{bmatrix} \pi_d \int (1-\alpha) A_t^{-\alpha} z^{1-\alpha} k^\alpha \ell_t^{\text{exit}}(s)^\nu \Phi_t(s) ds \\ \quad + (1 - \pi_d) \int (1 + \widetilde{\lambda}_t(s)) \left[(1-\alpha) A_t^{-\alpha} z^{1-\alpha} k^\alpha \ell_t^{\text{cont}}(s)^\nu - z i_t(s) \right] \Phi_t(s) ds \end{bmatrix}$$

$$\Omega_{t+1} = -\frac{1 - \pi_d}{1 + \widetilde{\Lambda}_{Ct} C_t^{-1} - \widetilde{\Lambda}_{Nt} w_t C_t^{-1}} \int \widetilde{\lambda}_t(s) b_t'(s) \Phi_t(s) ds, \text{ where } \Omega_{t+1} \text{ is from} \tag{68}$$

$$\Omega_t = \Lambda_{Ct} + \frac{1}{1 + r_t} \Omega_{t+1} \text{ starting from } \Omega_0 = 0$$

$$\widetilde{\Lambda}_{Nt} = \frac{1 - \pi_d}{1 + \widetilde{\Lambda}_{Ct} C_t^{-1} - \widetilde{\Lambda}_{Lt} w_t C_t^{-1}} \int \widetilde{\lambda}_t(s) \ell_t^{\text{cont}}(s) \Phi_t(s) ds \tag{69}$$