NBER WORKING PAPER SERIES

CAPITAL, IDEAS, AND THE COSTS OF FINANCIAL FRICTIONS

Pablo Ottonello
Thomas Winberry

Working Paper 32056
http://www.nber.org/papers/w32056

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2024

We thank Andy Abel, Fernando Alvarez, Cristina Arellano, David Argente, Marco Bassetto, Gideon Bornstein, Jeff Buckley, Satyajit Chatterjee, Nicolas Crouzet, Maarten De Ridder, Nik Engbom, Vincent Glode, Jeremy Greenwood, Hugo Hopenhayn, Chad Jones, Greg Kaplan, Pete Klenow, Pablo Kurlat, Simon Mongey, Guillermo Ordonez, Victor Rios-Rull, Richard Rogerson, Pierre-Daniel Sarte, Chris Tonetti, and seminar participants at the Jackson Hole Macro Conference 2023, the Penn-Wharton Macro Lunch Group, the Minneapolis Fed, the University of Wisconsin, the Richmond Fed, the conference in honor of Richard Rogerson's 65th birthday, the London School of Economics, the Philadelphia Fed, the Cowles Macro Conference 2023, the NBER Summer Institute 2023, the Econometric Society 2023 Australasia Meeting, Duke University, University of Surrey, Georgetown University, Cornell University, the University of Chicago, and Stanford University. We thank Samuel Stern, Jose Cristi Le-Fort, Erin Gibson, and Mariana Sans for excellent research assistance. Previously circulated as "Investment, Innovation, and Financial Frictions." The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Capital, Ideas, and the Costs of Financial Frictions
Pablo Ottonello and Thomas Winberry
NBER Working Paper No. 32056
January 2024
JEL No. E22,E23,G3,O30,O4

ABSTRACT

We study the role of financial frictions in determining the allocation of investment and innovation. Empirically, we find that firms are investment-intensive when they have low net worth but become innovation-intensive as they accumulate more net worth. To interpret these findings, we develop an endogenous growth model with heterogeneous firms and financial frictions. In our model, low net worth firms are investment-intensive because their returns to capital are high. Financial frictions slow the rate at which firms exhaust the returns to capital and shift towards innovation. Calibrating to the US economy, we find that the resulting lower growth implies large GDP losses even though capital misallocation is small. In other words, financial markets effectively fund the implementation of existing ideas, but do not adequately fund the discovery of new ideas. If innovation has positive spillovers, a planner would not only raise innovation but also lower investment expenditures among constrained firms.

Pablo Ottonello
Department of Economics
University of Maryland
Tydings Hall
7343 Preinkert Dr.
College Park, MD 20742
and NBER
ottonell@umd.edu

Thomas Winberry
University of Pennsylvania
2300 Steinberg-Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104
and NBER
twinb@wharton.upenn.edu

1 Introduction

In the long run, economic growth is driven by new ideas which push out the technological frontier. While many of these ideas are generated by new firms entering the economy, a substantial share of new ideas come from existing firms already in the economy. We study the innovation decisions of these existing firms because they face a novel tradeoff. Namely, because they already have ideas in place, existing firms must decide not only how much to innovate — creating new ideas — but also how much to invest — scaling up production using existing ideas. To the extent that a firm is financially constrained, these two activities will compete for the same funds within the firm. How do financial frictions distort the mix of investment and innovation within firms at the micro level? Do these distortions quantitatively matter for economic growth at the macro level?

We address these questions using new cross-sectional evidence and an endogenous growth model with heterogeneous firms subject to financial frictions. Empirically, we find a pecking order of firm growth: Firms are investment-intensive when they are small and have low net worth, but become more innovation-intensive as they grow and accumulate more net worth. Our model matches this finding because small firms have a high return to capital which crowds out innovation expenditures. At the micro level, financial frictions slow the rate at which firms accumulate capital, drive down its return, and shift into becoming innovation-intensive. At the macro level, these forces lower aggregate output by misallocating capital across firms with different productivities and by depressing innovation activities that improve the distribution of productivity. We calibrate the model to the US economy and find that the aggregate output losses from lower innovation are large even though the losses from capital misallocation are small. To the extent that innovation has positive spillovers, this allocation is not constrained efficient; a planner would not only raise innovation but also lower investment expenditures among constrained firms. Subsidies to innovation or investment do not fully achieve this goal because they do not generate the correct distribution of investment and innovation across firms.

Our empirical work focuses on how the composition of investment and innovation changes within firms as they grow and accumulate net worth over time. Our baseline sample is drawn from Compustat, a panel of publicly-listed U.S. firms. The long panel dimension of Compustat allows us to account for fixed differences across firms, which turns out to be important for our results. Of course, Compustat is a selected subsample of larger and older firms in the economy. We address this selection in our model by calibrating it to a broader set of firms in the economy and explicitly modeling the selection of firms into Compustat. We also find that our results hold in Orbis data, which contains some large private firms.

We document two key patterns about how firms allocate resources between investment and innovation in this data. First, firms become *less* investment-intensive as they accumulate net worth in the sense that their physical investment rates decline. Second, firms become *more* innovation-intensive as they accumulate net worth in the sense that their R&D expenditures and patenting rates increase. We infer that firms with a high shadow price of external finance, proxied by having low net worth, face a high return to capital but low return to innovation. We find similar patterns using other proxies for that shadow price, such as size — firms are investment-intensive when they are small but innovation-intensive when they are large — and age — firms are investment-intensive when they are young but innovation-intensive when they are old.

Motivated by this evidence, we develop a heterogeneous-firm endogenous growth model in which firms face financial constraints. In order to focus on the decisions of incumbent firms, we assume that new entrants simply draw an idea, embodied in their productivity, from the existing stock of ideas in the economy. Firms must then decide how much resources to spend on investment, which increases the capital stock used in production, and how much to spend on innovation, which increases the probability of receiving a new idea and raising the firm's productivity. The firm's mix of investment and innovation is determined by the relative return on these two activities. The return to capital is its marginal product and value as collateral in external borrowing, while the return to innovation is the probability of generating a new idea times the present value of that idea to the firm.

Our model generates the same pecking order of firm growth we found in the data. The return to capital is high for firms with low net worth because the marginal product of capital has not yet diminished and its use as collateral is more highly valued. As the firm accumulates capital and grows, the marginal product and collateral value of capital fall, inducing the firm to shift toward innovation. Financial frictions are central to this process because they control how quickly the firm can accumulate capital and shift into innovation.

The quantitative strength of the model's pecking order is governed not only by the financial frictions but also the innovation technology, which determines the return to innovation. Inferring the properties of this technology is difficult because its output, new ideas, are difficult to directly measure in the data. We infer the realization of new ideas using what firms reveal to us through their forward-looking investment decisions. In particular, our model predicts that new ideas should generate investment spikes — short-lived bursts of investment — in order to implement the new ideas in production. Consistent with this prediction, past R&D expenditures are strongly associated with investment spikes in our data; having one standard deviation higher R&D increases the probability of a spike in the following year by approximately 40%.

We calibrate our model to match this finding as well as other features of firms' investment, R&D, and borrowing in the US economy. We then validate our calibrated model using new empirical evidence on the response of innovation to changes in the after-tax price of investment induced by the Bonus Depreciation Allowance. In the data, we find that lower investment taxes significantly raise R&D expenditures, especially for small firms. We show that our model roughly matches the magnitudes of these regression coefficients, building confidence in its quantitative predictions.

We then use our calibrated model to quantify the aggregate effects of financial frictions. To do so, we compare the balanced growth path (BGP) in our model to a frictionless version of the model in which firms face no financial constraints. The frictionless model has no pecking order of firm growth because firms immediately accumulate their optimal scale of capital and begin innovating. Financial frictions generate two distortions relative to this benchmark. First, given the existing distribution of productivity, financial frictions misallocate capital across firms, lowering the *level* of aggregate TFP as in Hsieh and Klenow (2009). Second, financial frictions delay the rate at which firms begin innovating, lowering the *growth rate* of aggregate TFP.

Quantitatively, we find that financial frictions lower the aggregate growth rate by more than 40 basis points per year, leading to substantial output losses when cumulated over long horizons. For example, lower growth reduces GDP by nearly 25% over fifty years. In contrast, the output losses from capital misallocation are less than 5%. We show that this result holds for a range of parameter values governing the innovation technology, financial frictions, innovation spillovers, and firm entry. Hence, we conclude that the primary costs of financial frictions to the US economy are not that good ideas lack funding (capital misallocation), but that fewer good ideas are discovered in the first place (lower innovation).[1]

To the extent that innovation generates positive spillovers, this equilibrium allocation is socially inefficient, opening the door to policy intervention. In order to understand the nature of the externality, we first study the allocation chosen by a planner who internalizes the positive spillovers but is subject to the same financial constraints as individual firms. Clearly, the planner wants higher innovation, so we focus on the subtler question of how the planner reallocates investment in order to support this goal. The answer is heterogeneous across firms: higher innovation requires lower investment expenditures for constrained firms (due to the flow-of-funds constraint) but incentivizes higher investment for unconstrained firms (due to the complementarity of productivity and capital). We study how the planner balances these forces along a transition path starting from the equilibrium BGP. In the early phase of this transition, the substitutability for constrained firms dominates in the sense that aggregate investment falls. Over time, however, 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.

Unfortunately, simple policies like subsidies to innovation or investment cannot decentralize the planner's allocation because the planner's incentives vary across firms and over time. Nevertheless, we study the extent to which simple subsidies can partially achieve the planner's goals. First, we compute the effects of an innovation subsidy that increases aggregate innovation expenditures by the same amount as the social planner. The innovation subsidy raises the aggregate growth rate by 10% less than the planner because these expenditures are not optimally distributed across firms. Second, we compute the effects of an investment

[1]Because financial frictions primarily lower innovation by small firms, the majority of innovation activity in our model is concentrated among the largest, most unconstrained firms in the economy. We show that this distribution of innovation implies that financial crises, modeled as an unexpected tightening of financial constraints, do not persistently lower growth in the medium run.

tax cut of comparable size to the Tax Cuts and Jobs Act of 2017. We find that the tax cut also raises productivity growth because a higher capital stock also incentivizes innovation. This result contrasts with the neoclassical growth model, in which investment tax cuts have no effect on long-run growth.

Related Literature Our findings contribute to our understanding of the aggregate costs of financial frictions. The existing quantitative macro literature about financial frictions has primarily focused on how the frictions affect the allocation of inputs across firms (see, e.g., Hsieh and Klenow (2009), Buera, Kaboski and Shin (2011), Midrigan and Xu (2014), or Moll (2014)). However, these papers take the distribution of productivity as given, so the costs of financial frictions only arise from distorting capital as a function of productivity. In our model, financial frictions also distort the distribution of productivity by affecting innovation.

Our model combines elements of the Hopenhayn (1992) framework, in which firm dynamics are determined given an exogenous process for productivity, and the Schumpeterian growth framework 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 Schumpeterian models abstract from frictions to factor accumulation and, therefore, have no source of sluggish input dynamics.[2] We contribute to this literature by incorporating capital accumulation and sluggish dynamics induced by financial frictions.[3]

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 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), capital adjustment costs in Khan and Thomas (2008), or — most closely related to our model — financial frictions in Khan and Thomas (2013). We

[2]An exception is Bilal et al. (2021), who study a Schumpeterian model with labor search frictions.

[3]Our focus on differences in innovation intensity across firms is most closely related to Akcigit and Kerr (2018), who study how firms choose between two different types of innovation. We abstract from different types of innovation to instead study the choice between innovation and capital investment. In a related vein, Crouzet et al. (2022) develop a model in which firms choose between two different types of capital that differ in their degree of non rivalry. Our results are also related to the literature on how financial frictions distort the allocation of investment across different types of capital, such as durable vs. non-durable (Rampini, 2019), new vs. used (Lanteri and Rampini, 2023a), and clean vs. dirty (Lanteri and Rampini, 2023b).

contribute to this literature by incorporating innovation, which endogenizes the productivity process and therefore the distribution of optimal size.[4]

More broadly, our findings are consistent with a large body of empirical work that provides cross-country, regional, sectoral, firm-level, and case study evidence that better-functioning financial markets lead to higher economic growth (see Levine, 2005, for a detailed survey). There is also a large body of theoretical work about the relationship between financial markets and economic growth (see, e.g., Aghion, Howitt and Levine, 2018, and references therein). We contribute to this literature in at least two ways. First, we focus on how financial frictions distort firms' joint decisions between investment and innovation. Second, we focus on quantifying the aggregate effects of the resulting distortions.[5]

Road Map The rest of our paper is organized as follows. Section 2 begins with a word about the firms on which we focus our anlaysis. Section 3 documents the pecking order of firm growth in the data. Motivated by this evidence, Section 4 develops the model and Section 5 describes how the model matches the pecking order. Section 6 calibrates the key parameters of the model and shows the model matches a number of untargetted moments in the data. Section 7 uses the calibrated model to quantify the aggregate effects of financial frictions. Section 8 shows how innovation spillovers shape the constrained-efficient allocation and evaluates the effects of innovation subsidies and investment tax cuts. Section 9 concludes.

2 Which Firms Are We Thinking About?

Both our empirical and model analysis focus on a particular subset of firms in the economy. The vast majority of firms in the economy pursue little to no innovation and their 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 firms who eventually innovate and meaningfully

[4]Relatedly, Atkeson and Burstein (2010) embed innovation decisions in a Melitz (2003)-style model without capital in order to study the dynamic gains from trade. Chen and Xu (2023) incorporate both physical capital and R&D investments into an industry equilibrium model (in the tradition of Ericson and Pakes, 1995), abstracting from financial frictions.

[5]To our knowledge, existing quantitative work linking financial frictions and economic growth has abstracted from the firm-level trade-off between investment and innovation (see, for example, Queralto, 2020; Kalemli-Ozcan and Saffie, 2021, and references therein).

contribute to economic growth. We conceptualize the lifecycle of these firms in two phases. In the first *entry phase*, an entrepreneur uses their own time and skills to generate a new idea. If a new idea materializes, the entrepreneur creates a firm and implements the idea in practice. Once the production process using that idea is established, which may include the first couple years of the firm's life, the firm enters the *incumbent phase*. In this phase, the firm must decide how much to scale up its existing idea through investment and how much to attempt to generate a new idea through innovation.

We focus on the incumbent phase of these innovative firms for three main reasons. First, the incumbent phase contains the tradeoff between investment and innovation which motivates our paper. Second, the incumbent phase accounts for a significant share of aggregate R&D expenditures and patenting activity. Third, recent work by Akcigit and Kerr (2018) and Garcia-Macia, Hsieh and Klenow (2019) estimate that around three-quarters of aggregate growth comes from incumbent firms rather than new entrants.

As a result of focusing on incumbent firms, our analysis largely abstracts from how financial frictions affect innovation among new entrants. Therefore, our aggregate results should be interpreted as a lower bound on the growth effects of financial frictions.

3 Empirical Evidence

We show that firms are more investment-intensive when they have low net worth but become more innovation-intensive when they have high net worth.

3.1 Data Description

Our main analysis uses annual firm-level data from Compustat, a panel of publicly listed U.S. firms from 1975 – 2018. This dataset has two key advantages for our analysis. First, it contains information on firms' investment expenditures, R&D expenditures, and financial positions, allowing us to measure our key variables of interest. Second, it is a long panel, allowing us to absorb permanent differences across firms using fixed effects. To our knowledge, Compustat is the only US dataset with these properties.[6]

Of course, Compustat is a selected subsample of the types of innovative incumbent firms in which we are interested. We address selection into Compustat in two ways. First, we show that our results also hold in Orbis, which contains some larger private firms. Second, we calibrate the key features of our model to be representative of a broader set of firms in the economy, and explicitly account for the selection of firms into Compustat.

Our main variables of interest are firms' investment and innovation decisions as a function of their net worth. We measure the investment rate as the ratio of capital expenditures to the lagged value of plant, property, and equipment. Innovation activity is more difficult to measure, so we proxy for it in two ways. First, we proxy for the inputs into the innovation process using the *R&D share*, i.e. the ratio of R&D expenditures to the sum of R&D expenditures and capital investment. Second, we proxy for the outputs of the innovation process using approved patents collected from the United States Patent and Trademark Office by Kogan et al. (2017). We measure firms' net worth as the value of plant, property, and equipment, plus cash and short-term investments, minus total debt, consistent with our model. Appendix A describes the details of how we clean the data and presents descriptive statistics of our final sample. For our baseline analysis, we exclude observations associated with acquisitions in order to focus on innovation and investment occurring within firms (though we obtain similar results when including acquisitions).

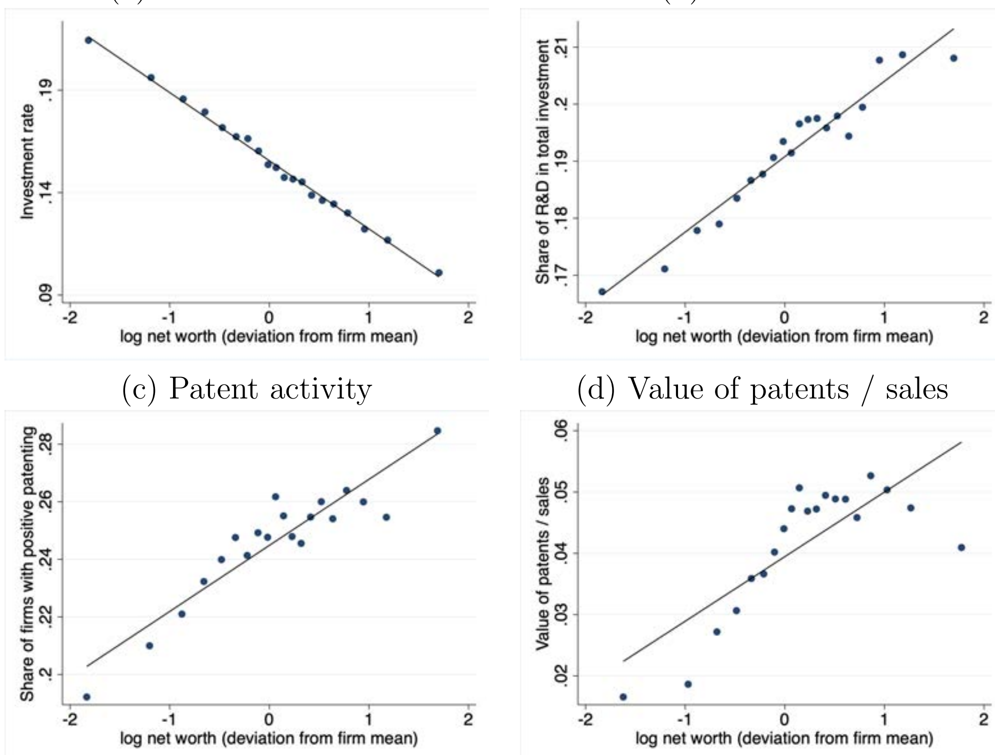
3.2 The Pecking Order of Firm Growth

We illustrate our pecking order of firm growth using simple binned scatterplots of investment and innovation by net worth. We focus on net worth because it is the state variable in our model which determines the firm's shadow price of external finance. We isolate within-firm variation by de-meaning both net worth and the outcome variables at the firm level, which is equivalent to using a firm fixed effect in a regression context. We condition on firms

[6] An alternative source of data on the investment and innovation decisions of privately-held firms is the Annual Survey of Manufacturers (ASM) from the U.S. Census. However, for small firms, these data are a rotating sample, which limits the within-firm variation we could use (see the discussion in Kehrig, 2015). In addition, these data only cover the manufacturing sector, while Compustat contains broader sectoral coverage of the economy.

Figure 1: The Pecking Order of Firm Growth

(a) Investment rates (b) R&D share



(c) Patent activity (d) Value of patents / sales

Notes: Binned scatter plots of investment rates, the R&D share, the share of firms with positive patenting, and the patent-value-to-sales ratio by the log of firm net worth. All variables are demeaned at the firm level. In order to make the units of the outcome variables more interpretable, we add back in the unconditional mean across all firms. For variable definitions and sample selection, see Appendix A.

with at least 20 years of observations in order to precisely estimate the firm-level mean, but Appendix A shows our results also hold for the whole sample of firms. In order to make the units of the outcome variables more interpretable, we add back in their mean values across all firms (which is a normalization that does not affect any results).

Figure 1 illustrates our two key empirical results. First, panel (a) shows that firms' investment activity *decreases* as they accumulate net worth; the smallest firms' investment rates exceed 0.2 but fall below 0.1 as they grow. This pattern is consistent with the notion that firms face a higher relative return to capital when they have low net worth.

Second, panels (b) – (d) shows that firms' innovation activity instead *increases* as firms accumulate net worth. In terms of innovation inputs, panel (b) shows that the R&D share increases by about 25% as firms grow. Appendix A shows that other measures of R&D activity, such as R&D-to-sales or the share of firms with positive R&D expenditures, are also increasing in net worth.[7]

In terms of innovation outputs, panels (c) and (d) show that patenting activity also increases as firms accumulate net worth. On the extensive margin, panel (c) shows that firms are about 30% less likely to obtain a successful patent when they are have low net worth compared to when they have high net worth. On the intensive margin, panel (d) shows that the market value of those new patents (scaled by firms' sales) doubles as firms grow.[8] Appendix A shows that other measures of patenting activity, such as the number of new patents per employee or the value of each new patent, are also increasing with net worth. While individually none of these measures of R&D or patenting activity fully captures firms' innovation, they are collectively consistent with the notion that firms face a higher relative return to innovation as they accumulate net worth.

Pecking Order by Size and Age Although net worth maps directly into the shadow price of external finance in our model, the corporate finance literature often uses various measures of size and age to proxy for that shadow price as well. We now show that our pecking order holds using these alternative sorting variables.

Table 1 shows the pecking order of firm growth holds for measures firm size. We summarize the pecking order using the regression

$$o_{jt} = \alpha_j + \gamma \log s_{jt} + \epsilon_{jt}, \tag{1}$$

where o_{jt} is the outcome of interest (investment rate, R&D share, indicator for positive patenting, or patent-value-to-sales) s_{jt} is the measure of size (net worth, capital, sales, or employment), and ϵ_{jt} is a residual. The coefficient of interest is γ, which measures how the outcome of interest varies with the particular measure of firm size. We standardize each size variable $\log s_{jt}$ over the entire sample in order to make the units of the coefficient γ easier

[7] A potential concern is R&D expenditures are under-reported in the data. We address this concern in Appendix A by conditioning on observations after the firm reports its first positive R&D expenditure and therefore has presumably set up the accounting infrastructure to report R&D expenditures going forward. We find similar results in this subsample.

[8] We use Kogan et al. (2017)'s measure of the market value of these patents, i.e. the change in firm equity value in a narrow window around the patent approval. We sum all of these changes that occur within a year to obtain the annual change in the firm's equity value due to new patents.

TABLE 1
THE PECKING ORDER BY VARIOUS MEASURES OF SIZE

	(1) Investment rate	(2) R&D share	(3) Patent activity	(4) Patent-value -to-sales
log net worth				
$\hat{\gamma}$	-0.068	0.024	0.049	0.021
	(0.003)	(0.003)	(0.007)	(0.005)
N	45935	47286	49105	31176
R^2	0.263	0.857	0.639	0.678
log capital				
$\hat{\gamma}$	-0.087	0.026	0.064	0.022
	(0.003)	(0.003)	(0.008)	(0.005)
N	49986	51569	53625	33561
R^2	0.272	0.852	0.633	0.671
log capital including intangibles				
$\hat{\gamma}$	-0.087	0.039	0.046	0.020
	(0.003)	(0.004)	(0.009)	(0.005)
N	44471	44794	46484	30549
R^2	0.282	0.852	0.634	0.670
log employment				
$\hat{\gamma}$	-0.045	0.002	0.110	0.008
	(0.005)	(0.005)	(0.011)	(0.006)
N	45050	45827	47508	30868
R^2	0.207	0.851	0.636	0.673
log sales				
$\hat{\gamma}$	-0.058	0.018	0.094	0.026
	(0.004)	(0.004)	(0.010)	(0.005)
N	49986	51569	53625	33561
R^2	0.218	0.851	0.634	0.671
Mean	0.13	0.16	0.34	0.05

Notes: Results from estimating the regression $o_{jt} = \alpha_j + \gamma \log s_{jt} + \epsilon_{jt}$, where o_{jt} is the outcome of interest (investment rate, R&D share, indicator for positive patenting, or patent-value-to-sales ratio; s_{jt} is the measure of size (net worth, capital, capital including intangibles, sales, employment); and α_j is a firm fixed effect. We standardize the size measures $\log s_{jt}$ over the entire sample. Standard errors, reported in parentheses, are clustered at the firm level. The variable "capital including intangibles" is from Peters and Taylor (2017) and is measured by incorporating both the firm's past investment and R&D expenditures. For variable definitions and sample selection, see Appendix A.

to interpret. We cluster standard errors at the firm level.

The first row of Table 1 quantifies the magnitudes and statistical significance of the bin-scatters using the regression (1) with s_{jt} = net worth. Column (1) shows that having one standard deviation more net worth than the average firm lowers the firm's investment rate by nearly 7 percentage points relative to the unconditional mean of 13 percentage points — a more than 50% decline in investment as firms accumulate net worth. Columns (2) – (4) show that having more net worth systematically raises our various proxies of innovation activity. For example, having one standard-deviation higher net worth increases the market value of that year's patenting activity relative to sales by 2 percentage points, a nearly 40% increase relative to its unconditional mean. All these effects are highly statistically significant.

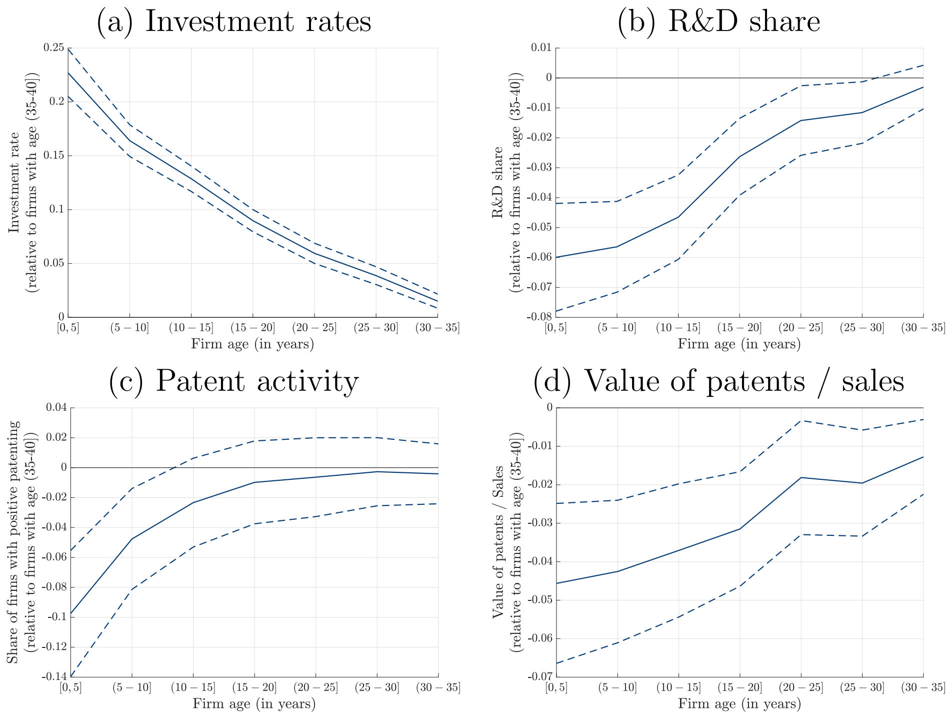
The remaining rows of Table 1 show that these patterns hold for the other measures of size. The second and third row proxy for size using physical capital (measured with plant, property, and equipment) or the sum of physical and intangible capital (as measured by Peters and Taylor (2017), which includes physical investment and R&D expenditures). The fourth and fifth rows proxy for size using employment or sales, which are common in the literature. The magnitudes of the pecking order are similar for all of these measures.

Figure 2 shows the pecking order by firm's age. We use Datastream to obtain age since incorporation (not age since IPO). We estimate the regression

$$o_{jt} = \alpha_j + \sum_{s \in \mathcal{S}} \gamma_s \text{age}_{sjt} + \epsilon_{jt}, \tag{2}$$

where age_{sjt} is a dummy variable that takes the value of 1 if the firm's age of incorporation in period t is within age bin $s \in \mathcal{S} \equiv \{[0,5],(5,10],(10,15],(15,20],(25,30],(30,35]\}$. We estimate the regression for firms with an age up to 40 years—above which the number of observations per bin becomes too small—so the omitted group corresponds to the age range $(35,40]$. Consistent with our main results, we find that firms are more investment-intensive when they are young but become more innovation-intensive as they become older.

Figure 2: The Pecking Order by Age



Notes: Results from estimating the regression $o_{jt} = \alpha_j + \sum_{s \in \mathcal{S}} \gamma_s \text{age}_{sjt} + \varepsilon_{jt}$, where o_{jt} is the outcome of interest (investment rate, R&D share, indicator for positive patenting, or patent-value-to-sales); age_{sjt} is a dummy variable that takes the value of 1 if firm's age of incorporation in period t is in group s and zero otherwise; and α_j is a firm fixed effect. We consider the following age groups: $\mathcal{S} \equiv \{[0,5], (5,10], (10,15], (15,20], (25,30], (30,35]\}$. We estimate the regression for firms with age up to 40 years, so the omitted group corresponds to age $(35, 40]$. Standard errors are clustered at the firm level and dashed lines correspond to 90% error bands. Firms' age since incorporation is obtained from Datastream. For other variable definitions and sample selection, see Appendix A.

3.3 Additional Results

Appendix A contains two sets of additional empirical results. First, we study our pecking order for other innovation outcomes, such as the R&D-to-sales ratio or patents-to-employee ratio, which have been studied in the existing literature. For the R&D-to-sales ratio, we find that the pecking order continues to hold in our sample, but only if we include firm fixed effects. We include fixed effects throughout our analysis to absorb fixed differences across firms that are absent in our model but drive a significant share of variation in the data. For the patents-to-employee ratio, we show that the pecking order holds for net worth and our measures of size other than employment. Hence, our results are consistent with Akcigit and Kerr (2018), who finds that patents-to-employees increases with employment using high-quality Census microdata. However, we focus on the pecking order as a function of net worth because it maps directly into our model of financial frictions.

Appendix A also contains a number of robustness checks of our main results. First, we show that the pecking order of firm growth also holds among Orbis data, which contains privately-held firms. Second, we show our results hold for other measures of innovation activity. Third, we show the pecking order also holds using within-sector, rather than within-firm, variation. Fourth, we show our results are robust to including time fixed effects to capture secular trends. Finally, we show our results are robust to including all firms in the sample rather than only those who have survived at least 20 years.

4 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.

4.1 Environment

We purposefully keep the entry phase of our model as simple as possible in order to focus on the tradeoff between investment and innovation in the incumbent phase.

Entry Phase There is a fixed flow π_d of new entrants each period that are endowed with zero debt and draw their initial levels of productivity and capital from some distribution $\Phi_t^0(z, k)$.[9] In order to capture imitation by new entrants (as in, e.g., Luttmer (2007)), we assume this distribution of productivity for new entrants is related to the distribution of productivity among incumbents; we will parameterize this dependence in Section 6. Imitation is necessary to ensure the model generates a positive growth rate along the balanced growth path (BGP). The equilibrium BGP of our model is identical to a version of the model with free entry in which the entry cost grows along with the economy, which Klenow and Li (2022) argue is the case empirically.[10]

[9] The initial capital stock can be interpreted as the equity stake injected by the household.

[10] However, neither our model nor the alternative with free entry would capture how financial frictions constrain the ability of potential entrants to come up their first idea. Quantitatively modeling this margin would require introducing a separate set of features, such as the innovation technology of new entrants, the ability of entrepreneurs to smooth consumption, and features of the venture capital market such as expert

Incumbent Phase Firms in the incumbent phase produce an undifferentiated good $y_{jt} = A_t z_{jt} k_{jt}^{\alpha}$ where j indexes a firm, z_{jt} is firm-specific productivity, k_{jt} is the firm's capital stock, and A_t is aggregate productivity (described below). Decreasing returns to capital $\alpha < 1$ ensure there is an optimal scale of the firm for each level of productivity, as in Hopenhayn (1992).[11,12] At the beginning each 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 debt. This exit shock occurs with probability π_d, which is also the inflow of new entrants, ensuring the total mass of firms in production is constant over time.[13] The exit shocks ensures that firms do not outgrow their financial frictions in the long run.

Firms that will continue into the next period spend resources on 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 i_{jt} increases the probability of realizing a successful innovation, $\eta(i_{jt})$. We assume the arrival probability $\eta(i_{jt})$ is increasing, concave, and bounded between 0 and 1. A successful innovation permanently raises productivity by a factor Δ:

$$\log z_{jt+1} = \left\{ \begin{array}{r} \log z_{jt} + \Delta + \varepsilon_{jt+1} \text{ with probability } \eta(i_{jt}) \\ \log z_{jt} + \varepsilon_{jt+1} \text{ with probability } 1 - \eta(i_{jt}) \end{array} \right\}, \tag{3}$$

where $\varepsilon_{jt+1} \sim N(0, \sigma_\varepsilon)$ are idiosyncratic shocks to productivity growth unrelated to innovation.[14] Successful innovations can capture the arrival of new technologies, production

advice and the dilution/moral hazard of founders. We prefer to abstract from these issues and focus on how financial frictions affect incumbent firms. Therefore, we interpret our analysis as a lower bound on the effect of financial frictions on long-run growth.

[11] Appendix B shows that this production function can be derived from a model in which labor is a variable input in production: $y_{jt} = A_t z_{jt} k_{jt}^{\widetilde{\alpha}} \ell_{jt}^{\widetilde{\nu}}$ with $\widetilde{\alpha} + \widetilde{\nu} < 1$. The production function in the main text is equal to the variable profit function $\max_{\ell_{jt}} A_t z_{jt} k_{jt}^{\widetilde{\alpha}} \ell_{jt}^{\widetilde{\nu}} - w_t \ell_{jt}$. In this case, the production parameter α reflects the elasticity of revenue with respect this combination of inputs through $\alpha = \frac{\widetilde{\alpha}}{1-\widetilde{\nu}}$. We calibrate the model under this interpretation. Hence, incorporating variable labor would not affect our results, so we omit it from the main text for the sake of parsimony.

[12] 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 interpret decreasing returns as reflecting a downward-sloping demand curve or if we allow for free entry.

[13] A fixed mass of firms is the only specification that yields a stable average firm size in the model with endogenous labor in Footnote 11 under the assumption of fixed population growth.

[14] These idiosyncratic shocks generate realistic dispersion across firms in the data, which are useful in

practices, or new products. Following the Schumpeterian growth literature, we assume the arrival of one of these events induces a discrete jump in productivity.

The total amount of resources required to achieve success probability $\eta(i_{jt})$ is $i_{jt} \times (A_t z_{jt})^{\frac{1}{1-\alpha}}$. This cost specification has two natural properties. First, the fact that it grows along with aggregate productivity ensures the model has a BGP. Second, the fact that it scales with individual productivity implies that all financially unconstrained firms have the same growth rate, a property known as Gibrat's law. This property provides useful benchmark both because Gibrat's law arguably holds among large firms in the data and because it is a common feature of Schumpeterian models. In contrast, constrained firms grow faster than unconstrained firms in our model because constrained firms have a higher marginal product of capital.

Firms cannot sell their existing ideas, i.e. innovation expenditures must be non-negative $i_{jt} \geq 0$. In principle, financially constrained firms may have an incentive to sell their ideas in order to finance investment. In practice, the "market for ideas" — licensing arrangements, patent sales, mergers and acquisitions, etc. — is rife with frictions. We view our assumption that $i_{jt} \geq 0$ as the limit in which those frictions are sufficiently large to prevent trade in the market for ideas altogether. These frictions allow the model to generate inaction in innovation rates, which is common in the data.[15]

Firms have two sources of finance for their investment and innovation expenditures. First, they can borrow externally, but this borrowing is subject to the collateral constraint $b_{jt+1} \leq \theta k_{jt+1}$. 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. Second, firms can use their internal resources, but they cannot raise new equity. This assumption implies that dividend payments must be nonnegative:

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

This no equity-issuance constraint may seem overly strong at face value in light of the role of venture capital play in financing young firms in the technology sector. However,

replicating empirical regressions on model-simulated data.

[15] In our Compustat sample, nearly 50% of firm-year observations have zero recorded R&D expenditures.

equity issuance among established incumbent firms, which are the focus of our model, is relatively rare. Section 5 also shows that our pecking order also holds if ideas are partially collateralizable, as would be the case with equity issuance or earnings-based constraints.

Aggregate productivity A_t captures the positive spillovers from one firms' innovation decisions onto others:

$$A_t = \left(\int z_{jt} dj \right)^a, \tag{4}$$

where $a \geq 0$ governs the degree of spillovers, e.g. the degree to which others' ideas are relevant or can be appropriated for use. We choose this form of spillover to cleanly illustrate how financial frictions interacts with the positive externality.

Key Differences Between Capital and Ideas While we have made specific parametric choices in our model, there are three important economic differences between capital k and ideas z which drive our results:

(i) *Technological*: capital is subject to decreasing returns to scale, so all else equal the marginal product of capital is higher for small firms. This force leads to a higher return to capital for these firms.

(ii) *Collateral value*: capital is more collateralizable than ideas in external borrowing, further increasing the return to capital for small firms.

(iii) *Resale frictions*: ideas are difficult to sell in the secondary market, leading to inaction in innovation expenditures.

Household To close the model, 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

$$\frac{1}{1+r_t} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma}. \tag{5}$$

4.2 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 $n_{jt} = A_t z_{jt} k_{jt}^{\alpha} + (1-\delta)k_{jt} - b_{jt}$.[16] Exiting firms set $k_{jt+1} = b_{jt+1} = i_{jt} = 0$, while continuing firms' decisions are characterized by the Bellman equation

$$v_t^{\text{cont}}(z,n) = \max_{k',i,b'} n - k' - (A_t z)^{\frac{1}{1-\alpha}} i + \frac{b'}{1+r_t} + \frac{1}{1+r_t}\mathbb{E}_t\left[v_{t+1}(z',n')\right] \text{ s.t. } d \geq 0 \text{ and } b' \leq \theta k', \quad (6)$$

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 $b_t'(z,n)$; distribution 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 the Bellman equation (6); (ii) the evolution of $\Phi_t(z,n)$ is consistent with firm decisions; (iii) the real interest rate r_t is given by (5) with $C_t = \int \left(y_{jt} - (k_{jt+1} - (1-\delta)k_{jt}) - (A_t z_{jt})^{\frac{1}{1-\alpha}} i_{jt}\right) dj$; and (iv) aggregate productivity is given by the definition (4).

Balanced Growth Path Because productivity grows over time, the limiting behavior of the model exhibits a balanced growth path. Along the BGP, all macroeconomic aggregates grow at the same rate $1 + g = (1+\widetilde{g})^{\frac{1+a}{1-\alpha}}$, where $\widetilde{g}$ is the growth rate of mean firm-specific productivity $\mathbb{E}_t[z_{jt}]$.[17] Appendix B provides details.

[16] In fact, given that productivity has a unit root, it may be possible to further reduce the firm's state variable to the ratio n_{jt}/z_{jt}. We do not pursue this option for expositional convenience.

[17] Our model is a fully endogenous growth model in the sense that the extended version with labor described in Footnote 11 exhibits what Jones (2005) terms "strong scale effects." However, we could eliminate these strong scale effects by assuming that new ideas increase productivity $\log z_{jt}$ by $\Delta \times (A_t \mathbb{E}_t[z_{jt}])^{\phi}$ with $\phi < 0$ (our baseline model corresponds to $\phi = 0$). As discussed by Jones (2005), the effects of various counterfactuals or policies at any point along a transition path are continuous in the parameter ϕ.

5 The Pecking Order of Firm Growth

We now show that our model generates a pecking order of firm growth consistent with the data. We discuss the key economic forces governing this pecking order, motivating how we calibrate the model in Section 6. We also show that our results are robust to two natural extensions of the model.

5.1 Characterizing Decision Rules

In order to characterize the firm's decision rules, we first note that the marginal cost of spending resources on either investment or innovation 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 resources inside the firm and is therefore the opportunity cost of instead spending those resources on investment or innovation. Appendix B shows two straightforward results about the shadow value of funds. First, the shadow value is equal to $\frac{\partial v_t(z,n)}{\partial n} = 1 + \lambda_t(z,n)$, where $\lambda_t(z,n)$ is the Lagrange multiplier on the non-negativity constraint on dividends; that is, the shadow value of funds is equal to the household's value, 1, plus the shadow price of issuing equity, $\lambda_t(z,n)$. The second result is that firms equate this shadow price $\lambda_t(z,n)$ to the shadow price of additional borrowing when constrained, i.e. the expected value of the multipliers on future collateral constraints $\mu_t(z,n)$ in all possible states of the world. Hence, the multiplier $\lambda_t(z,n)$ encodes how both financial frictions affect the marginal cost of firm growth through either investment or innovation. We refer to the multiplier $\lambda_t(z,n)$ as the *financial wedge*.

Building on this discussion, Proposition 1 characterizes firms' optimal choices of investment and innovation. This proposition extends a similar result from Khan and Thomas (2013)'s 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 such that*

(i) ***Financially unconstrained****: If $n \geq \overline{n}_t(z)$, then the financial wedge $\lambda_t(z,n) = 0$. Being financially unconstrained is an absorbing state. The capital accumulation $k_t^*(z)$, innovation $i_t^*(z)$, and borrowing $b_t^*(z)$ policies are independent of net worth.*

(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****: If $n \in (\underline{n}_t(z), \overline{n}_t(z))$, the collateral constraint is not currently binding $b' < \theta k'$ but the financial wedge is positive $\lambda_t(z,n) > 0$.*

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

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

$$1 + \lambda_t(z,n) = \frac{1}{1+r_t}\mathbb{E}_t\left[(MPK_{t+1}(z',k') + 1 - \delta) \times (1 + (1-\pi_d)\lambda_{t+1}(z',n'))\right] + \theta\mu_t(z,n) \tag{8}$$

$$\begin{aligned} 1 + \lambda_t(z,n) &\geq \frac{\eta'(i)}{(A_t z)^{\frac{1}{1-\alpha}}}\frac{1}{1+r_t}\left(\mathbb{E}_t[v_{t+1}(z',n'|\iota_{t+1}(z,n)=1] - \mathbb{E}_t[v_{t+1}(z',n'|\iota_{t+1}(z,n)=0)]\right), \\ \textit{with} &= \textit{ if } i_t(z,n) > 0 \end{aligned} \tag{9}$$

where $MPK_{t+1}(z',k') = \alpha A_{t+1} z'(k')^{\alpha-1}$ is the marginal product of capital, $\lambda_t(z,n)$ is the Lagrange multiplier on the no equity issuance constraint $d \geq 0$, $\mu_t(z,n)$ is the multiplier on the collateral constraint $b' \leq \theta k'$, and $\iota_{t+1}(z,n)$ denotes the realization of a successful innovation.

Proof. See Appendix B. ∎

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 and are indifferent over any combination of external financing b' and internal financing d leaves them financially unconstrained. Following Khan and Thomas (2013), we resolve this indeterminacy by requiring that firms pursue the "minimum savings policy," i.e., the smallest level of b' that leaves them unconstrained with probability one (see Appendix B).

The remaining firms are affected by financial frictions in some way. *Currently constrained* firms' collateral constraint binds in the current period, directly limiting their ability to borrow. *Potentially constrained* firms do not face a binding collateral constraint in the current period, but may reach 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 investment and innovation decisions for any of these three types of firms. Equation (7) is the nonnegativity constraint on dividends, which binds if 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 resources or new borrowing. Equations (8) and (9) are the first-order conditions for investment and innovation.[18]

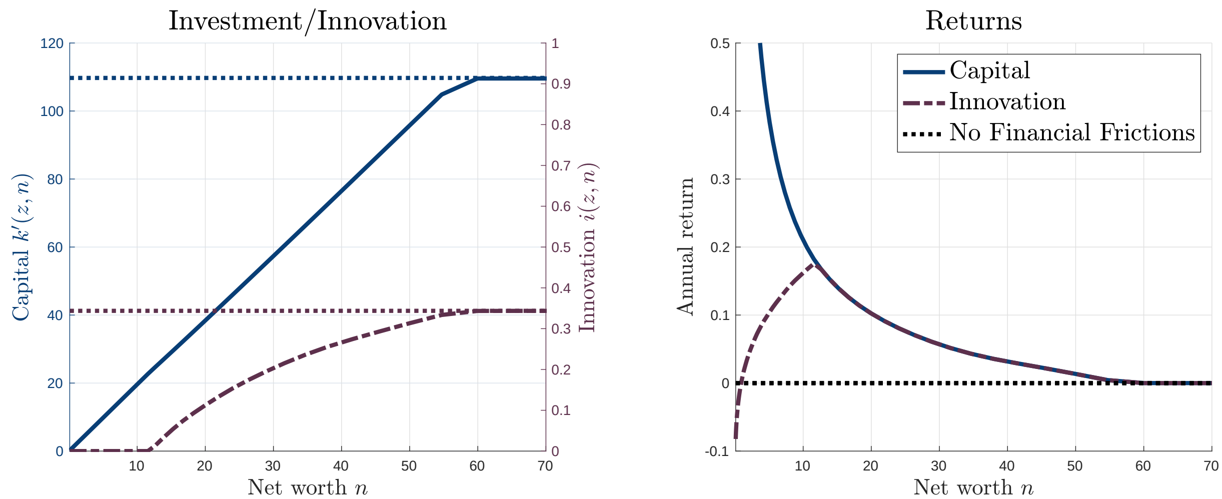
As discussed above, the marginal cost of investment or innovation on the LHS of (8) and (9) is the shadow value of funds $1 + \lambda_t(z, n)$. The marginal benefit on the right-hand side of the investment first-order condition (8) is given by two terms: the discounted expected marginal product of capital 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, the innovation first-order condition (9) may not hold with equality because firms face a non-negativity constraint on innovation $i_t(z, n) \geq 0$. The marginal benefit of innovation on the right-hand side of (9) is the marginal improvement in the probability of success per unit of innovation expenditure times the expected increase in firm value from a successful innovation.

5.2 Illustrating the Pecking Order

We illustrate the model's pecking order by plotting the decision rules which emerge from these first-order conditions. We generate these plots using our calibrated parameter values from Section 6, but the qualitative properties that emerge hold for a wide range of the parameter space. The left panel of Figure 3 plots the investment and innovation policies $k'_t(z, n)$ and $i_t(z, n)$ as a function of net worth, holding fixed the level of productivity z. The right panel plots the net returns associated with each activity, i.e. the right-hand side of the respective first-order condition minus one.

[18] 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 (7) - (9). 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 endogenous 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. See Appendix C for details.

FIGURE 3: The Pecking Order of Firm Growth in the Model



Notes: the left panel plots capital expenditures $k_t'(z,n)$ (left axis) and innovation intensity $i_t(z,n)$ (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 (8) and (9) 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.

The model's pecking order can be summarized by three distinct regions in net worth space. First, for the lowest levels of net worth,, the firm spends all of its available resources on capital and sets innovation expenditures $i_t(z,n) = 0$ because the return to capital lies strictly above the return to innovation. As the firm accumulates capital, it drives down the return capital due to the diminishing marginal product and the lower shadow value of collateral. At the same time, higher capital also raises the return to innovation because TFP and capital are complements in production. However, since the firm does not innovate in this region, it only grows its size by accumulating capital.[19]

As the firm continues to grow and accumulate net worth, the returns to capital and innovation eventually intersect and the firm begins innovating. In this second region of the pecking order, the innovation first-order condition (9) holds with equality, so the returns to capital and innovation must be equalized. However, both (net) returns are still strictly greater than zero because the financial wedge is positive $\lambda_t(z,n) > 0$, implying that firms do not pay dividends. In this case, investment and innovation are substitutes because higher

[19]Appendix D shows that most of the gap between the return to capital and innovation is due to the higher marginal product of capital, not its collateral value.

investment must be accompanied by lower innovation for a given level of net worth (or vice-versa). On the other hand, an increase in net worth will increase both investment and innovation, with the associated sensitivities determined by the slope of the expenditure curves. Hence, in this region of the pecking order, the firm grows both through capital accumulation and through the potential realization of successful innovations.

For sufficiently high levels of net worth, the firm enters the last region of the pecking order in which it is near its optimal scale given its current level of productivity, $k_t^*(z)$. At this point, the net returns to investment and innovation are close to zero, implying that the firm's policies become independent of net worth. In this case, 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 and the firm may re-enter an earlier region of the pecking order.

This discussion illustrates how our model is consistent with the empirical pecking order that we documented in Section 3. First, firms enter the economy with a new idea but less capital than the implied optimal scale, $k < k_t^*(z)$.[20] This initial condition places most new entrants in the first region of the pecking order in which they start growing only through investment. Second, established firms who receive a new idea from a successful innovation will similarly enter a situation in which their current capital stock is below their new, higher optimal scale $k < k_t^*(z)$. These firms will again enter an earlier region of the pecking order and prioritize investment before innovating again.

This discussion also shows that our model matches two stylized facts about scale-dependent firm growth in the data. First, employment and sales of small firms grow faster than large firms (see, e.g., Akcigit and Kerr, 2018); this occurs in our model because small firms have a high marginal product of capital in the early stages of the pecking order. Second, the average growth rates of large firms are the same because all unconstrained firms choose the same innovation intensity $i_t(z, n)$ and therefore have the same probability of receiving a new idea (see the plots in Appendix D).[21]

[20] In this sense, firms enter the economy rich in ideas but poor in capital. This assumption is consistent with empirical evidence on firm's life-cycle growth patterns (e.g., Haltiwanger, Jarmin and Miranda, 2013).

[21] The fact that Gibrat's law does not hold for all firms in our model complicates aggregation relative to the typical endogenous growth model; specifically, we need to keep track of the entire distribution of firms in order to solve the model (see Appendix C for details).

5.3 Role of Key Parameters

We have found that two sets of parameters are quantitatively important in shaping the pecking order: the degree of financial frictions and the efficiency of the innovation technology.

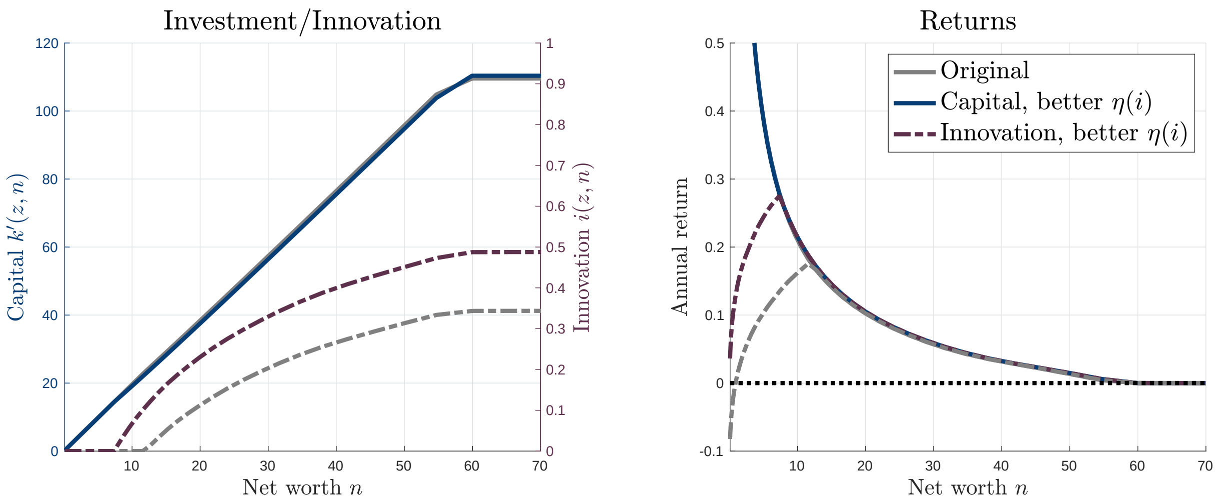
Role of Financial Frictions Figure 3 shows that, without financial frictions, the model would not have a pecking order in the first place; instead, firms would immediately lever up to their optimal scale given current productivity. In this case, investment and innovation would become independent of net worth, size, and age, inconsistent with the evidence presented in Section 3. In this sense, financial frictions are the key model ingredient which allows it to be consistent with the pecking order of firm growth.

While some form of financial constraints are necessary to generate the pecking order in our model, the precise form we've chosen is not. In general, financial constraints play two roles in our model. First, they imply that small firms have a higher shadow value of funds $1 + \lambda_t(z, n)$ and, therefore, face a higher marginal cost spending resources. Any financial constraint with this feature will imply that small firms face a high marginal product of capital, generating a high return to capital. Second, the financial constraints determine the collateral value of either investment or innovation, which affects the return on either activity. In our model, the financial constraint $b_{jt+1} \leq \theta k_{jt+1}$ implies that capital is collateralizable but ideas are not. However, Section 5.4 below extends the model to allow for collateralizable ideas and shows that the pecking order still holds in this extension.

Other sources of adjustment frictions, such as capital adjustment costs, customer base accumulation, or learning how to implement new ideas may also match pecking order of firm growth. We focus on financial frictions given the empirical importance of net worth. Financial frictions uniquely imply that investment and innovation are substitutes for constrained firms, which will be important for our policy analysis and allow the model to match the size-dependent response of innovation to investment tax shocks.

Role of Innovation Technology Figure 4 illustrates the effects of a more efficient innovation technology, which raises the success probability $\eta(i)$ for any level of innovation intensity i. The higher success probability shifts up the returns to innovation, which implies

Figure 4: Role of Innovation Technology in the Pecking Order



Notes: the left panel plots capital expenditures $k'_t(z,n)$ (left axis) and innovation intensity $i_t(z,n)$ (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 (8) and (9) minus 1. "Better $\eta(i)$" refers to the model with higher η_0 than in our baseline calibration (see Section 6).

that it intersects the returns to capital at a lower level of net worth. Therefore, firms begin innovating earlier on, and conditional on innovating, do more innovation.

5.4 Extensions

We show that the pecking order holds in two natural extensions of the model: allowing ideas to be partially collateralizable in external borrowing and allowing for heterogeneity in the size of successful innovations.

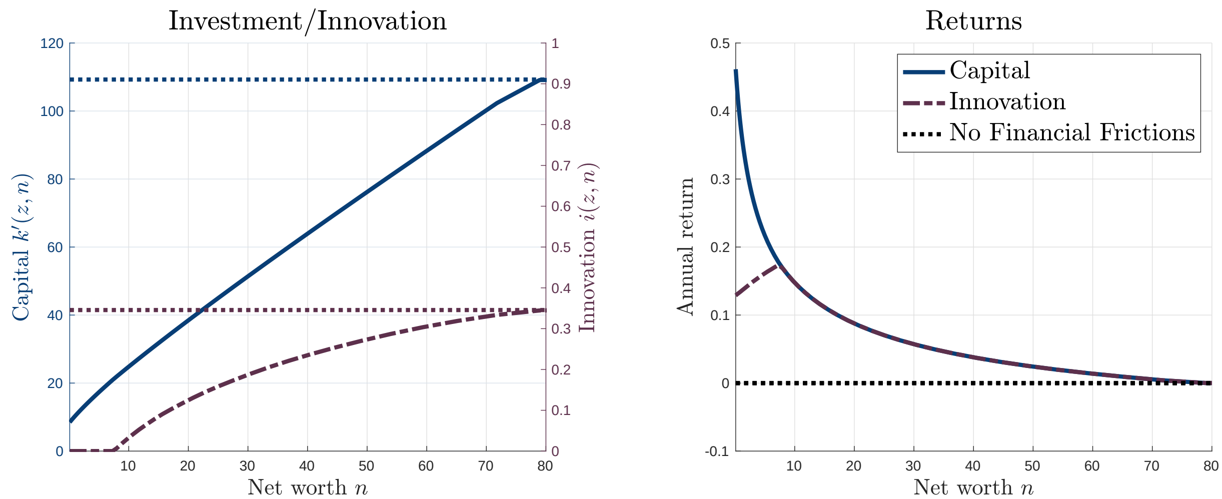
Collateralizable Ideas We consider the alternative borrowing constraint[22]

$$b_{jt+1} \leq \widetilde{\theta}\mathbb{E}_t\left[A_{t+1}z_{jt+1}k^{\alpha}_{jt+1}\right]. \tag{10}$$

This specification captures the spirit of the earnings-based constraints documented by Lian and Ma (2021), Greenwald et al. (2019), and Caglio, Darst and Kalemli-Özcan (2021). We

[22]This functional form requires that the idiosyncratic productivity shocks ε_{jt+1} are bounded in order to support any borrowing. In practice, we bound the shocks in the interval $[-3\sigma_\varepsilon, 3\sigma_\varepsilon]$.

Figure 5: The Pecking Order with Collateralizable Ideas



Notes: the left panel plots capital expenditures $k_t(z,n)$ (left axis) and innovation intensity $i_t(z,n)$ (right axis) in market equilibrium BGP of the alternative model with the financial constraint (10). The right panel plots the return to investment and innovation from the alternative first-order conditions

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

$$\begin{aligned}1+\lambda_t(z,n) \geq\ & \frac{\eta'(i)}{(A_tz)^{\frac{1}{1-\alpha}}}\frac{1}{1+r_t}\left(\mathbb{E}_t[v_{t+1}(z',n'|\iota_{t+1}(z,n)=1]-\mathbb{E}_t[v_{t+1}(z',n'|\iota_{t+1}(z,n)=0)]\right)\\ &+\frac{\eta'(i)}{(A_tz)^{\frac{1}{1-\alpha}}}\left(\mathbb{E}_t[A_{t+1}z'|\iota_{t+1}(z,n)=1]-\mathbb{E}_t[A_{t+1}z'|\iota_{t+1}(z,n)=0)]\right)(k')^{\alpha}.\end{aligned}$$

"No financial frictions" refers to the model in which all firms following the unconstrained policies $k^*(z)$ and $i^*(z)$ but using the same real interest rate from the market BGP.

assume expected future earnings enter this constraint in order to allow for future ideas, not just current ideas, to be partially collateralizable.[23] In this sense, the alternative constraint also has similar features to equity issuance. We recalibrate the strength of the alternative constraint $\widetilde{\theta}$ to match the same leverage as in our baseline model; therefore, the key difference between this alternative constraint and our baseline constraint is the fact that future ideas are now collateralizable.

Figure 5 shows that this alternative constrained generates a similar pecking order to our baseline model. The primary quantitative difference is that the return to innovation is shifted up relative to the baseline model, reflecting the fact that future ideas are now partially

[23]This property amplifies the positive externality of innovation relative to our baseline model because new ideas directly shift out everyone's borrowing constraint (the right-hand side of (10) is increasing in A_{t+1}).

collateralizable. But, overall, the pecking order is robust to allowing for the collateralizability of ideas. We will maintain with our baseline model of financial frictions $b_{jt+1} \leq \theta k_{jt+1}$ for the rest of the paper because it is canonical in macro-finance (e.g. Kiyotaki and Moore, 1997).

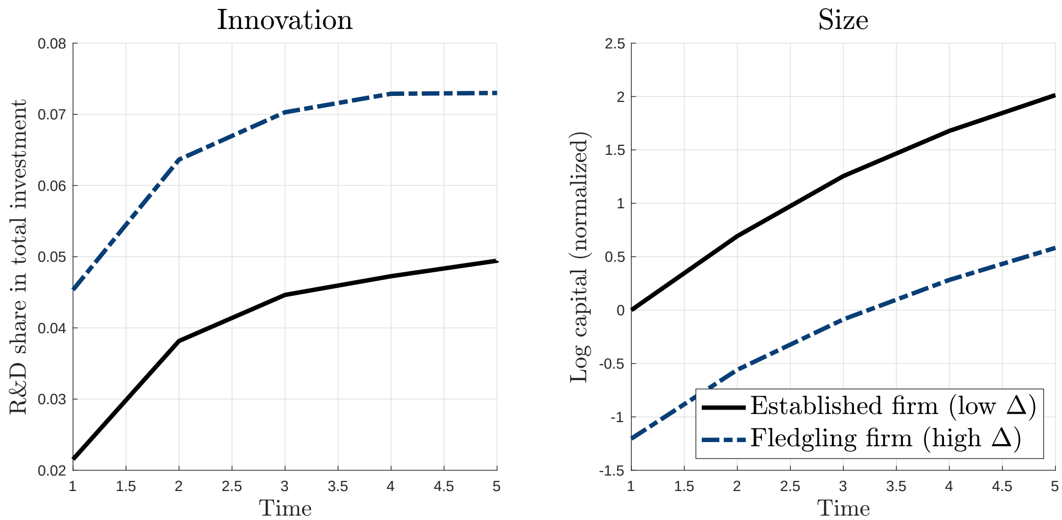
Heterogeneity in Innovation Size In Section 3 and Appendix A.2, we discussed how some studies argue that small firms in the economy are more innovation-intensive than larger firms. While we did not find strong evidence for this result in our sample, we cannot rule out the possibility that the smallest and youngest firms in the economy are not more innovation-intensive before they enter our sample. For example, these new firms may have a comparative advantage in innovation relative to other firms, leading them to do more innovation. In this extension, we show that our model can accommodate this possibility without affecting our pecking order of growth within firms.

Specifically, we compare the time paths of two firms that differ in the size of their successful innovations, Δ. The first firm is representative of a small incumbent in our Compustat sample; it has our calibrated innovation technology and starts with a level of capital similar to a smaller firm within Compustat (we discuss how we model the selection into Compustat in Section 6). The second firm is representative of a brand new firm with a higher return to innovation than incumbents; its innovation technology has a larger increase in productivity upon success Δ, but the firm starts with a lower level of capital. We then simulate the time paths of the R&D share and the capital stock for these two firms over five years.

Figure 6 shows that innovation is negatively correlated with size *across firms*, as some in the existing literature argue, but innovation is still positively correlated with size *within firms*, as in our pecking order of firm growth. The across-firm correlation is negative because the smaller firm has a higher technological return to innovation Δ, and therefore pursues more innovation and less investment for any level of net worth. However, the within-firm correlation is still positive because each firm is more willing and able to finance higher innovation as they accumulate net worth. Given this robustness, we abstract from heterogeneity in the technological return to innovation Δ for the rest of the paper.[24]

[24] Akcigit and Kerr (2018) develop a model without capital but with two types of innovation activities that differ in the size of their success, loosely mapping into heterogeneity in our Δ parameter. In their model, innovation declines in firm size because firms endogenously choose to do the lower-Δ type of innovation as

Figure 6: The Pecking Order with Heterogeneity in Innovation Size



Notes: figure plots time path of R&D-to-sales ratios (left panel) and detrended log capital stock (right panel) for two types of firms: one with the baseline innovation technology, and one with a better technology with higher step size Δ but starts with a smaller log capital stock. Log capital in right panel is normalized to zero in the first year for the firm with the baseline technology.

6 Parameterization

We now calibrate the model to ensure it is in line with key features of the data. Section 6.1 describes the moments we use to discipline the key parameters governing the pecking order described above. Section 6.2 uses these moments, and others, to calibrate the model. Finally, Section 6.3 shows that the calibrated model matches various untargeted statistics, including the response of innovation to investment tax shocks.

---

they grow. Akcigit and Kerr (2018) discipline the strength of this force using the fact that small firms' employment grows faster than large firms. Our model also implies that small firms grow faster than large firms, but because small firms have a higher marginal product of capital rather than a better innovation technology. We therefore view our mechanism and Akcigit and Kerr (2018)'s mechanism as complementary channels which may both be operative depending on the type of firm being studied. However, only our mechanism predicts that firm-level innovation rates increase in firm size, as we found in the data.

6.1 Strategy for Disciplining Key Forces

Following much of the literature, we will choose the tightness of the collateral constraint θ to match the average leverage of firms in the data.[25] Therefore, the main challenge in our calibration is to pin down the properties of the innovation technology, i.e., the probability of a successful innovation $\eta(i)$ and the size of successful innovations Δ.

While we can arguably measure innovation inputs using R&D expenditures, there is no direct measure of the output, successful innovations. Given this difficulty, we instead infer successful innovations through what firms reveal through their forward-looking investment decisions. In our model, firms that receive a successful innovation experience an *investment spike*—a large but short-lived surge in their investment rate — in order to adapt their capital stock to the new, higher level of productivity.[26] Therefore, the responsiveness of investment spikes to R&D expenditures should be informative about the innovation technology.

We study the relationship between investment spikes and R&D expenditures in our Compustat data. Following Cooper and Haltiwanger (2006), we define investment spikes as years in which a firm's investment rate is above 20%. In our sample, the frequency of investment spikes is 23% and the average size of an investment spike is 37%, similar to Cooper and Haltiwanger (2006)'s Census sample.

We estimate the linear probability model

$$\mathbb{1}\{\frac{x_{jt}}{k_{jt}} \geq 0.2\} = \alpha_j + \alpha_{st} + \sum_{h=1}^{H} \beta_h \left(\frac{\text{RD}_{jt-h}}{\tilde{y}_{jt-h}}\right) + \Gamma' X_{jt} + \epsilon_{jt}, \tag{11}$$

where $\frac{x_{jt}}{k_{jt}}$ denotes the investment rate of firm j in period t; $\frac{\text{RD}_{jt}}{\tilde{y}_{jt}}$ the R&D-to-sales ratio; α_j and α_{st} firm and time by 4-digit sector fixed effects; X_{jt} is a vector of firm-level controls; and ϵ_{jt} is a residual. Our coefficient of interest, β_1, measures how the probability of an investment spike is related to previous R&D expenditures. The vector X_{jt} includes variables that attempt to control for two alternative reasons for investment spikes that are unrelated

[25]While our model abstracts from some determinants of firm leverage, like the tax advantage of debt, we show that our calibration implies realistic responses of investment and innovation to investment tax shocks.

[26]This approach relies on the assumption that a successful innovation creates a discrete jump in productivity while shocks unrelated to productivity, ε_{jt}, follow a normal distribution. Again, our formulation is aimed at capturing the breakthrough nature of innovation, which are classic elements in Schumpeterian growth models (e.g., Aghion and Howitt, 1992; Grossman and Helpman, 1991).

TABLE 2
R&D EXPENDITURES PREDICT INVESTMENT SPIKES

	(1)	(2)	(3)
$\frac{\text{RD}_{jt-1}}{\tilde{y}_{jt-1}}$	1.27	1.09	1.10
	(0.16)	(0.15)	(0.15)
$\frac{\text{cf}_{jt}}{k_{jt}}$		0.11	0.11
		(0.03)	(0.03)
years since spike$_{t-1}$			0.003
			(0.0008)
$\frac{k_{jt}}{n_{jt-1}}$			-0.050
			(0.008)
Observations	55,647	55,647	55,647
Adj. R^2	0.280	0.297	0.300

Notes: Results from estimating $\mathbb{1}\{\frac{x_{jt}}{k_{jt}} \geq 0.2\} = \alpha_j + \alpha_{ts} + \sum_{h=1}^{4} \beta_h \left(\frac{\text{RD}_{jt-h}}{\tilde{y}_{jt-h}}\right) + \Gamma' X_{jt} + \epsilon_{jt}$, where $\frac{x_{jt}}{k_{jt}}$ denotes the investment rate of firm j in period t; $\frac{\text{RD}_{jt}}{\tilde{y}_{jt}}$ 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 residual. 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 A.

to the innovation technology. First, we include 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. Second, we control for the number of years since a previous investment spikes and the capital-to-labor ratio, which are informative about the effects of non-convex capital adjustment costs.[27] Appendix A.1 details the construction of these additional control variables. We set $H = 4$ for our baseline model and explore alternative lags in robustness analysis. We two-way cluster standard errors by firm and year.

Table 2 shows that, consistent with the predictions of our model, R&D expenditures are a strong predictor of investment spikes. Column (1) reports the estimated coefficient β_1 from the linear probability model (11) without any additional controls X_{jt}. Quantitatively, the estimated coefficient implies that having last year's an R&D-to-sales ratio one standard

[27] (S, s) models predict that an investment spike is more likely when the firms' capital stock k is farther from its optimal scale $k^*(z)$. This "gap" is increasing if firms have not had an investment spike in the recent past, or if the capital-to-labor ratio is far from normal (under the assumption that the choice of labor is more flexible and therefore better reflects current productivity z).

deviation above the mean increases the probability of an investment spike by 7 percentage points, i.e. a 30% increase in the probability of a spike relative to its unconditional mean. Column (2) shows that this estimate survives controlling for changes in cash flow which may independently affect both investment and R&D expenditures.

One may be concerned that investment spikes in the data are driven by non-convex capital adjustment costs, not the arrival of new ideas as in our model. We address this concern in three ways. First, we target the passthrough of R&D expenditures to investment spikes, not the overall frequency of investment spikes. Second, fixed costs and irreversibilities most naturally occur at the unit of the plant or even production line, and Compustat firms aggregate over many such units. Finally, Column (3) in Table 2 shows that our regression coefficient is virtually unaffected by controlling for the years since the last spike and the capital to labor ratio.

While we admittedly 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 A.4 presents robustness analysis and additional supportive evidence about the relationship between R&D expenditures and investment spikes. We show that the results presented in Table 2 are robust to using an alternative definition of investment spikes that considers a sector-level threshold instead of an absolute threshold, using more or less lags of R&D-to-sales ratios, and using additional controls used in the investment literature (e.g., size, sales growth, and the share of current assets). We also present a complementary event study which shows that R&D-to-sales tend to increase before investment spikes.

6.2 Calibration

We calibrate the model in two steps. First, we fix a subset of parameters to match standard targets. Second, we choose the remaining parameters so that moments from the model's BGP match key features of the data, including the regression coefficients documented above.

For the calibrated model, we assume that productivity is drawn from log-normal distribution whose mean z_0 equals the mean of the distribution of incumbent firms, and set the dispersion in those draws to $\sigma_z = \Delta$. We further assume that all new entrants start with an initial capital stock k_0 roughly equal to 4% of the average capital stock in the economy. This value is between the two possibilities discussed by Khan and Thomas (2013): having new entrants' capital be 10% of average capital or having their employment (in the extended model with labor discussed in Footnote 11) be 10% of average employment. We show that our main results are robust to alternative choices for this initial distribution in Section 7.[28]

Fixed Parameters Table 3 contains the parameters that we exogenously fix. We set the EIS $1/\sigma = 1.5$, in line with estimates from the finance literature. We make this choice because changes in the real interest rate are very powerful in our model given that firms face no other adjustment costs. We set the EIS on the high end of estimates from the literature to dampen this unrealistic interest-sensitivity of investment. Given this value of the EIS, we set the household's discount factor β so that the real interest rate is 4% annually along the BGP. We set the elasticity of output with respect to inputs to be $\alpha = 0.55$, close to the 0.59 value Cooper and Haltiwanger (2006) estimate for manufacturing plants.[29] We set 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.

Fitted Parameters Table 4 contains the endogenously chosen parameters and the moments that we target in the data. The first three targets and the standard deviation of investment rates $\sigma(x_{jt}/k_{jt})$ are drawn from our Compustat data. In our model, we mirror the sample selection into Compustat by conditioning on firms that have survived at least five years. This choice matches the median time to IPO of seven years from Ottonello and Winberry (2020), but allowing the "entry phase" outside our model to take two years. Hence, the key assumption we make are that the innovation technology and the determinants of

[28]While this entry process is representative for the typical firm, it may not capture "revolutionary entrants" who enter the with a substantially better idea than other firms in the economy. These revolutionary entrants can be thought of as coming from another distribution of initial state variables with very high productivity z. Financial frictions would be even more binding on these firms because their optimal scale of capital would be even further above their initial capital stock.

[29]In the model with labor discussed in Footnote 11, our choice of α is consistent with an underlying production function in which the labor share is 2/3 and the total returns to scale is 0.85.

Table 3
Fixed Parameters

Parameter	**Description**	**Value**
Household		
β	Discount factor	0.97
$1/\sigma$	EIS	1.50
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.

investment volatility are similar within and outside of Compustat conditional on age.

We now discuss the identification of the parameters in more detail. The first three parameters in the left panel govern the innovation technology. We assume that the probability of success is given by $\eta(i) = 1 - \exp\{-\eta_0 [(1+i)^{\eta_1} - 1]\}$, which is increasing, concave, and bounded between 0 and 1. In addition, $\eta(0)$ is finite, allowing the model to generate inaction in R&D expenditures as discussed earlier. Under this functional form, the parameter η_0 governs the overall efficiency of the success probability with respect to innovation intensity while the parameter η_1 governs the slope of this relationship. The parameter Δ then controls the size of successful innovations.

While all parameters are jointly chosen to match all moments, we have found that the innovation technology is primarily pinned down by the first three targets in the right panel of Table 4. The regression coefficient from Section 6.1 has a strong influence over both parameters that govern the probability of success, η_0 and η_1. In contrast, the average R&D to sales ratio, $\mathbb{E}[\text{RD}_{jt}/y_{jt}|\text{RD}_{jt} > 0]$, primarily influences the curvature parameter η_1 because it governs how quickly the marginal benefit of additional R&D spending is exhausted.[30] Given these two targets, the average size of investment spikes $\mathbb{E}[x_{jt}/k_{jt}|\text{spike}]$ pins down the size of successful innovations Δ.

The degree of financial frictions is parameterized by the collateral constraint θ. This parameter governs how much firms can borrow and is therefore primarily pinned down by

[30] We target the average R&D-to-sales ratio conditional on positive R&D because the R&D inaction rate in Compustat (45%) is much higher than in our model (10%). We prefer to not target this inaction rate because Compustat inaction may partly reflect misreporting rather than true inaction.

Table 4

Fitted Parameters and Empirical Targets

Parameter	**Description**	**Value**	**Target** (all joint)	**Data**	**Model**
Innovation technology					
η_0	Idea arrival function	1.01	Regression coefficient	1.09	1.03
η_1	Idea arrival function	0.21	$\mathbb{E}[\mathrm{RD}_{jt}/y_{jt} \vert \mathrm{RD}_{jt} > 0]$	0.06	0.06
Δ	Size of innovations	0.10	$\mathbb{E}[x_{jt}/k_{jt} \vert \text{spike}]$	0.37	0.35
Financial frictions					
θ	Collateral constraint	0.52	$\mathbb{E}[b_{jt}/k_{jt}]$	0.34	0.30
Productivity shocks					
σ_ε	SD of shocks	0.03	$\sigma(x_{jt}/k_{jt})$	0.15	0.13
Innovation spillovers					
a	Innovation spillovers	0.55	Growth rate	0.02	0.02

Notes: left panel contains the parameters chosen to match the moments in the right table. "Idea arrival function" refers to $\eta(i) = 1 - \exp\{-\eta_0 [(1+i)^{\eta_1} - 1]\}$. "Regression coefficient" is the regression coefficient β_1 from Table 2 column (2). $\mathbb{E}[\mathrm{RD}_{jt}/y_{jt}|\mathrm{RD}_{jt} > 0]$ is the average R&D to sales ratio for observations with positive R&D expenditures in the Compustat sample described; in the model, we compute R&D expenditures as $A_t z^{\frac{1}{1-\alpha}} i_t(z,n)$. $\mathbb{E}[x_{jt}/k_{jt}|\text{spike}]$ is the average size of investment spikes in the Compustat sample. $\mathbb{E}[b_{jt}/k_{jt}]$ is the average gross leverage of firms in the QFR reported by Crouzet and Mehrotra (2020). $\sigma(x_{jt}/k_{jt})$ is the standard deviation of investment rates in our sample. "Growth rate" is the aggregate growth rate along the BGP.

average leverage $\mathbb{E}[b_{jt}/k_{jt}]$. We target average gross leverage of firms in the Quarterly Financial Reports (QFR) from Crouzet and Mehrotra (2020). Importantly, this dataset contains private firms and is therefore more representative than our Compustat sample.

Given these targets, the dispersion of idiosyncratic shocks σ_ε helps us match the dispersion of investment rates in the data. The degree of innovation spillovers a is then inferred residually to match a long-run growth rate of 2% per year.

Table 4 shows that the model matches the targeted moments fairly well (even though it is overidentified due to nonlinearity). Importantly, the model is very close to the targets informative about both the innovation technology and the degree of financial frictions. Averaging across all firms, the implied excess return to capital is 5.4% and the excess return to innovation (conditional on innovating) is 2.7%.

The calibrated parameter values are broadly consistent with the range of estimates in the literature. The collateral constraint implies that about half of tangible capital is collateralizable. 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 Schumepterian 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 0.74. One interpretation of this finding is that investment spikes capture a broader set of innovations than do patents.

6.3 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 strength of financial frictions in driving investment and innovation decisions in our model.

Firm Heterogeneity Appendix D analyzes the two sources of firm heterogeneity, lifecycle dynamics and productivity differences (due to either successful innovations or productivity shocks). Following the pecking order of firm growth, 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 Appendix D also studies the response of innovation to changes in the after-tax price of investment induced by the Bonus Depreciation Allowance, a countercyclical investment stimulus used during the 2001 and 2008 recessions. Following Zwick and Mahon (2017), we exploit sectoral heterogeneity in the policy treatment to estimate the effects of the Bonus using a difference-in-difference empirical design. We first reproduce Zwick and Mahon (2017)'s finding that the Bonus substantially raises firm-level investment in our Compustat data. We then show that the Bonus also raises R&D expenditures, especially for small firms. We replicate this experiment in our model by feeding in a comparable shock to the relative price of investment and reproducing the empirical regression specifications. We find that the model roughly matches all of these regression coefficients. Hence, our model is quantitatively consistent with the cross-price elasticities of innovation to investment prices by firm size.

7 Aggregate Costs of Financial Frictions

We now use the calibrated model to compute the macroeconomic effects of financial frictions.

The Costs of Financial Frictions Financial frictions reduce aggregate output through two channels in our model. First, as described in Section 5, they reduce innovation expenditures for constrained firms. Aggregating across firms, lower innovation reduces the long-run growth rate through

$$g^* \approx \frac{1}{1-\alpha}(1+a)(e^{\Delta}-1)\int \iota_{jt}dj. \tag{12}$$

We quantify this channel by comparing our calibrated model to the *frictionless model* in which there are no financial frictions. In the frictionless model, all firms follow the unconstrained policies $k^*(z)$ and $i^*(z)$ from Proposition 1, which imply a higher arrival rate of new ideas ι_{jt}.

Second, financial frictions distort the allocation of capital across firms with different levels of productivity, reducing the level of TFP as in the misallocation literature. Formally decomposing the misallocation costs of financial frictions separately from the growth costs is conceptually difficult in our model because changes in the allocation of capital also affect innovation decisions. We instead provide a simple upper bound for the misallocation costs by computing the counterfactual level of output which solves

$$Y_t^* = \max_{k_{jt}} A_t \int z_{jt} k_{jt}^{\alpha} dj \quad \text{such that} \int k_{jt} dj \leq K_t.$$

This counterfactual holds aggregate capital fixed, but distributes it across firms to maximize current output given the current distribution of productivity. Appendix B shows that the

Table 5
Aggregate Output Losses from Financial Frictions

	Lost growth per year	20 years	Lost growth after 30 years	40 years	50 years	Misallocation costs (upper bound)
Baseline	41bps	8.6%	13.1%	17.9%	22.8%	5.0%
Higher η_0	54bps	11.0%	16.9%	23.2%	29.8%	4.8%
Higher η_1	54bps	11.0%	16.9%	23.2%	29.8%	4.6%
Higher θ	33bps	6.7%	10.2%	13.8%	17.5%	3.9%
Lower a	39bps	7.9%	12.1%	16.5%	21.1%	5.2%
Lower z_0	35bps	7.1%	10.8%	14.7%	18.7%	3.8%
Higher k_0	39bps	7.9%	12.1%	16.5%	21.0%	4.6%

Notes: output losses from financial frictions computed relative to frictionless model in which all firms follow the unconstrained policies $k^*(z)$ and $i^*(z)$ from Proposition 1. "Lost growth" is the difference in the BGP growth rate g^* in the frictionless model vs. the full model. "Lost growth after" cumulates the lost growth over different horizons. "Misallocation (upper bound)" refers to the losses from capital misallocation discussed in the main text. "Baseline" refers to calibration model. "Higher" and "lower" refer to sensitivity analysis with respect to parameter values 25% higher or lower than their calibrated value.

ratio of actual output to this optimal level is

$$\frac{Y_t}{Y_t^*} = \frac{\int z_{jt} \left(\frac{k_{jt}}{K_t}\right)^{\alpha} dj}{\left(\int z_{jt}^{\frac{1}{1-\alpha}} dj\right)^{1-\alpha}}.$$

Our upper bound on the misallocation costs of financial frictions is then $\frac{Y_t}{Y_t^*} - 1$.[31]

Quantitative Results The top row of Table 5 shows that the growth costs of financial frictions are large: without financial frictions, the long-run growth rate would be more than 40 basis points higher per year. Cumulated over long horizons, this differences implies substantial losses in aggregate output; for example, after fifty years, GDP in the frictionless model is nearly one-quarter higher than in the calibrated model. In contrast, the misallocation costs are fairly modest, with an upper bound of 5% of output.

In this sense, the main quantitative impact of financial frictions is that fewer new ideas are discovered, not that existing ideas go underfunded. This finding complements the macro-development literature which studies the costs of financial frictions in developing economies. For example, Buera, Kaboski and Shin (2011) argue that the misallocation costs of financial

[31]This object is an upper bound on misallocation costs in the sense that it assumes capital can be reallocated after the realization of the productivity shock z_{jt}, which firms cannot do in our model.

frictions can be up to 40% of aggregate TFP in developing economies.[32] In contrast, our model is calibrated to the US economy, in which financial markets are much better developed and therefore misallocation is small. We nevertheless find that in this context financial frictions are costly due to reduced innovation.

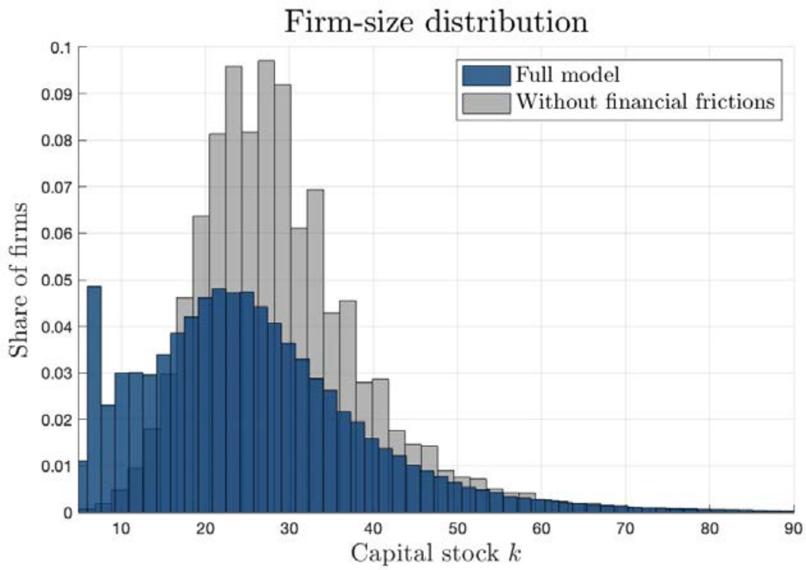
Sensitivity Analysis Table 5 also performs sensitivity analysis with respect to key parameters in the model. The first two rows show that making the innovation technology more efficient (raising η_0 or η_1 by 25%) increases the growth costs but decreases the misallocation costs of financial frictions. This occurs because higher η_0 or η_1 raise the marginal benefit of innovation, implying that more firms are constrained in terms of innovation relative to investment. The third row of Table 5 shows the effects of raising the collateral constraint by 25%, reducing the degree of financial frictions. As expected, less financial frictions reduce both the growth and misallocation costs. However, the growth costs are still substantial.

The remaining rows of Table 5 contain further sensitivity analysis. Lower innovation spillovers a reduces the growth effects of financial frictions, as expected from (12). Lower productivity among new entrants z_0 implies that entrants are closer to their efficient scale and therefore less constrained than in the baseline calibration, which reduces the growth and misallocation costs of financial frictions. Raising the capital stock of new entrants k_0 similarly implies that entrants start out closer to their efficient scale, reducing the growth and misallocation costs. In all cases, the growth costs are still substantial.

Distributional Effects of Financial Frictions Looking across firms, financial frictions depress innovation primarily in small, financially constrained firms for whom the return to capital is relatively high. This force thickens the right tail of the firm size distribution relative its mode. We illustrate this mechanism in Figure 7, which compares the distribution of detrended capital stocks in our calibrated model and the frictionless model. 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 dis-

[32] Of course, the quantitative magnitudes of misallocation are debated in the literature. For example, Midrigan and Xu (2014) argue that the TFP losses from misallocation are around 5-10% in the modern sector of Korea.

Figure 7: Distributional Effects of Financial Frictions



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 B). "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.

tribution 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.[33]

Because most innovation is done by financially unconstrained firms, Appendix D shows that temporary financial shocks θ_t do not have a particularly persistent effect on aggregate growth 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 economic effects in the data (e.g., Cerra and Saxena, 2008). Based on representative

[33] 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. 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.

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.[34]

8 Policy Implications

The equilibrium we've studied so far is not socially efficient because firms do not internalize the positive spillovers from their innovations. In order to better understand the implications of this externality, Section 8.1 studies a constrained-efficient planner who internalizes the externality. Section 8.2 uses these results to study how two commonly-used policies, innovation subsidies and investment tax cuts, address this externality

8.1 Planner's Allocation

We characterize the problem of a constrained-efficient social planner who faces the same financial constraints as private firms but internalizes the positive spillovers from innovation.[35] In principle, this planner may also want to change the private allocation due to pecuniary externalities through the real interest rate (which may affect welfare due to market incompleteness). We exclude them from the main text because they do not affect the long-run choices of the planner and have already been extensively studied in the literature (see, for example, Geanakoplos and Polemarchakis, 1986; Lorenzoni, 2008; Dávila and Korinek, 2018).[36]

Appendix B formulates the planner's problem recursively. The problem is technically challenging because the state variable is the entire distribution of firms and the control

[34]One caveat to this result is that we abstract from how financial frictions affect firm entry. Ates and Saffie (2021) study how financial shocks affect aggregate productivity through the composition of firms' entry in the context of a small open economy experiencing a sudden stop.

[35]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.

[36]It is straightforward to incorporate pecuniary externalities; results available upon request.

variables are entire functions of the firms' individual states. We overcome this challenge by solving the problem in the function space following Lucas and Moll (2014) and Nuño and Moll (2018). We arrive at the following characterization of the solution:

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

$$\omega_t^{cont}(z,n) = \max_{k',b',i} n - k' - (A_t z)^{\frac{1}{1-\alpha}} i + \frac{b'}{1+r_t} + \Lambda_t z + \frac{1}{1+r_t}\mathbb{E}_t[\omega_{t+1}(z',n')] \ s.t. \ d \geq 0 \ and \ b' \leq \theta k' \tag{13}$$

where Λ_t is the planner's shadow value of the innovation externality:

$$\Lambda_t = a\left(\int z_{jt}dj\right)^{a-1} \times \int (1+\lambda_{jt})\left[z_{jt}k_{jt}^{\alpha} - \frac{1}{1-\alpha}(A_t^{\frac{\alpha}{1-\alpha}} z_{jt})^{\frac{1}{1-\alpha}} i_{jt} dj\right] \tag{14}$$

with the convention that $\lambda_{jt} = i_{jt} = 0$ for exiting firms.

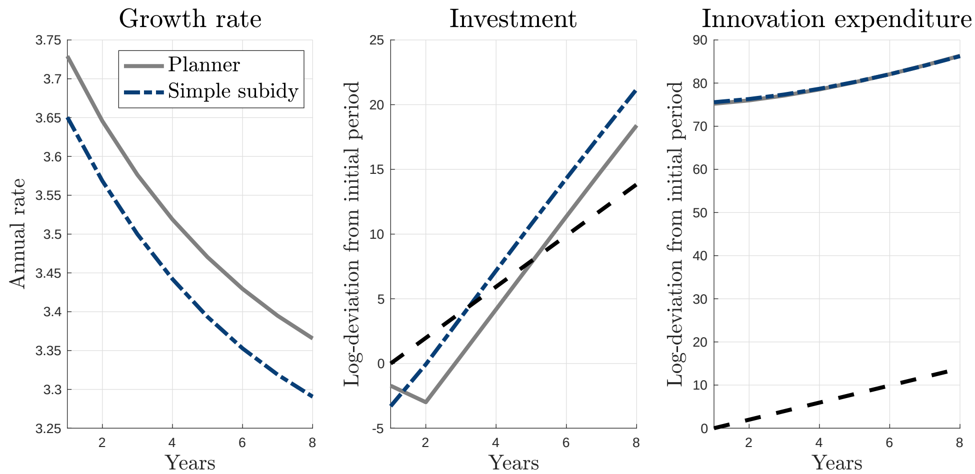
Proof. See Appendix B. ■

The only difference between the private Bellman equation (6) and the planner's augmented Bellman equation (13) is the shadow value of the innovation externality, Λ_t. Equation (14) 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.[37] This object is itself an integral of a product of two firm-level objects: the marginal increase in production net of innovation costs, $z_{jt}k_{jt}^{\alpha} - \frac{1}{1-\alpha}(A_t^{\frac{\alpha}{1-\alpha}} z_{jt})^{\frac{1}{1-\alpha}} i_{jt}$, times the firms' shadow value of funds, $1 + \lambda_{jt}$. Hence, financial frictions amplify the positive externality of innovation; higher production raises cash flows and alleviates financial constraints, which the planner values at the shadow price λ_{jt}.

While the planner clearly prefers more innovation than in equilibrium ($\Lambda_t > 0$), it faces a tradeoff in terms of investment. On the one hand, higher innovation for constrained firms requires less investment expenditures due to their flow-of-funds constraint, i.e. investment and innovation are substitutes for constrained firms. On the other hand, higher innovation

[37]Consistent with our focus on incumbent firms, 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.

Figure 8: Planner's Allocation and a Simple Innovation Subsidy



Notes: aggregate transition paths chosen by planner (grey lines) and generated by the simple 23% innovation subsidy (dashed blue lines). Growth rate in top rate is in percentage points per year. Aggregate 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.

incentivizes more investment for unconstrained firms due to the complementarity between productivity and capital in production, i.e. investment and innovation are complements for unconstrained firms. In order to characterize this tradeoff, we solve for the transition path chosen by the planner starting from the equilibrium BGP.

Figure 8 characterizes this tradeoff along the transition path chosen by the planner starting from the market equilibrium. Early on, the substitutability between investment and innovation dominates in the sense that aggregate investment falls. However, as the distribution of net worth improves over time, the complementarity between investment and innovation begins to dominate in the sense that aggregate investment eventually increases. Appendix D.5 contains additional analysis of the planner's allocation.

8.2 Evaluating Practical Policies

The planner's allocation seems difficult to implement in practice because one has to get both the allocation of innovation *and investment* correct and the relevant tradeoffs vary across both firms and time.[38] 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 Figure 8 compares the planner's aggregate allocation to a simple innovation subsidy that is uniform across firms and time. We choose the subsidy rate to generate the same increase in aggregate innovation expenditures chosen by the planner. However, the planner's allocation generates 10 basis points more growth per year than the comparable subsidy. This result occurs because the subsidy disproportionately increases innovation among unconstrained firms, who have a lower return to innovation expenditures $\eta'(i)$. Hence, the innovation subsidy does not fully replicate the planner's solution because it fails to deliver the correct distribution of innovation across firms.

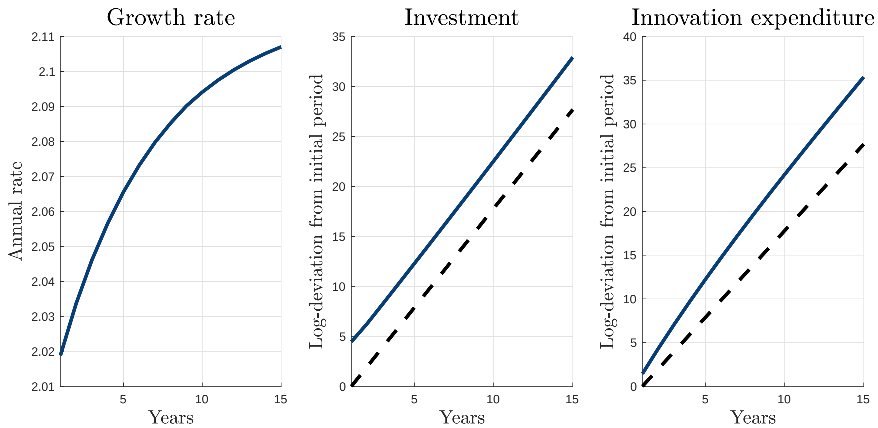
Investment Tax Cut We illustrate the connection between investment tax cuts and innovation using the Tax Cuts and Jobs Act (TCJA 2017) as an example. Appendix B.3 shows that the tax system changes the after-tax price of investment to $1 - \tau\zeta_t$, where τ is the corporate tax rate and ζ_t is the present value of tax deductions per unit of investment. The TCJA 2017 raised the present value of deductions to $\zeta_t = 1$, lowering the relative price of investment. We mirror this policy change in our model by studying a permanent decline in the after-tax price of investment of the same size.

Figure 9 shows that, in our model, full expensing increases the long-run growth rate by 10 basis points per year. 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. However, these positive effects take fifteen years to fully materialize.

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,

[38] The planner's augmented Bellman equation (13) 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 9: Investment Tax Cuts Raise Innovation (TCJA 2017)



Notes: transition path following an unexpected, permanent decline in the relative price of capital 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. Investment and innovation expenditures expressed as log-deviations from initial period.

only lead to an increase in the level of output (not its growth rate).

9 Conclusion

In this paper, we have studied the efficiency costs of financial frictions for the macroeconomy. While the quantitative macroeconomic literature has focused on how financial frictions distort investment decisions and misallocate capital, we focus on how financial frictions distort innovation and lower economic growth. We showed these two margins are empirically linked through the pecking order of firm growth. Quantitatively, we found that the US financial system effectively funds the implementation of existing ideas, but do not adequately fund the discovery of new ideas. A key contribution of our paper is a new endogenous growth framework with heterogeneous firms and financial frictions that is consistent with this evidence and can be used to draw aggregate implications.

We have purposefully kept our framework as parsimonious as possible in order to focus on the novel mechanisms for our research questions. However, the parsimony of our framework can be leveraged in order to obtain additional insights. For example, extending the model to include labor would also incorporate a negative pecuniary externality of innovation operating through the labor market. In this extension, innovations from unconstrained firms raise labor demand, which in turn raises labor costs and tightens financial constraints on affected firms. The optimal policy would have to balance the tradeoff between growth (coming from the innovation spillovers) and misallocation (coming from these pecuniary externalities).

Another important extension would relax our assumption that firms cannot sell ideas. Given the frictions in the market for ideas, it is natural to use a search-and-matching model in the spirit of Lucas and Moll (2014), Perla and Tonetti (2014), or Akcigit, Celik and Greenwood (2016). In this extension, firms would choose between spending some time producing output and the remaining time searching in the market for ideas. This extension would provide a third source of firm-level growth, technology adoption. We conjecture that low-productivity firms would be more likely to adopt than innovate because their time cost of searching is relatively low and their expected return from matching is relatively high. Conversely, high-productivity firms would be more likely to sell ideas, especially if they are financially constrained and would like to finance investment. This extension would also endogenize the innovation spillovers through the composition of idea trades that emerges in the market for ideas.

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, L., P. Howitt, and R. Levine (2018): "Financial development and innovation-led growth," in *Handbook of finance and development*, pp. 3–30. Edward Elgar Publishing.

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.

Ates, S. T., and F. E. Saffie (2021): "Fewer but better: Sudden stops, firm entry, and financial selection," *American Economic Journal: Macroeconomics*, 13(3), 304–356.

Atkeson, A., and A. T. Burstein (2010): "Innovation, firm dynamics, and international trade," *Journal of political economy*, 118(3), 433–484.

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.

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.

Caglio, C., M. Darst, and Kalemli-Özcan (2021): "Collateral Heterogeneity and Monetary Policy Transmission: Evidence from Loans to SMEs and Large Firms," *NBER working paper*, (w28685).

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.

Chen, Y., and D. Y. Xu (2023): "A structural empirical model of r&d, firm heterogeneity, and industry evolution," *The Journal of Industrial Economics*, 71(2), 323–353.

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

CROUZET, N., J. EBERLY, A. L. EISFELDT, AND D. PAPANIKOLAOU (2022): "Intangible capital, non-rivalry, and growth," *Available at SSRN*.

CROUZET, N., AND N. R. MEHROTRA (2020): "Small and large firms over the business cycle," *American Economic Review*, 110(11), 3549–3601.

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

DAVIS, S. J., J. C. HALTIWANGER, AND S. SCHUH (1998): "Job creation and destruction," *MIT Press Books*, 1.

ERICSON, R., AND A. PAKES (1995): "Markov-perfect industry dynamics: A framework for empirical work," *The Review of economic studies*, 62(1), 53–82.

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.

HALTIWANGER, J., R. S. JARMIN, AND J. MIRANDA (2013): "Who creates jobs? Small versus large versus young," *Review of Economics and Statistics*, 95(2), 347–361.

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.

HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 124(4), 1403–1448.

HURST, E., AND B. W. PUGSLEY (2011): "What do small businesses do?," Discussion paper, National Bureau of Economic Research.

Jones, C. I. (2005): "Growth and ideas," in *Handbook of economic growth*, vol. 1, pp. 1063–1111. Elsevier.

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

Kalemli-Ozcan, S., and F. Saffie (2021): "Finance and Growth: Firm Heterogeneity and Creative Destruction," .

Kehrig, M. (2015): "The cyclical nature of the productivity distribution," *Earlier version: US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15*.

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.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988): "Production, growth and business cycles: I. The basic neoclassical model," *Journal of monetary Economics*, 21(2-3), 195–232.

Kiyotaki, N., and J. Moore (1997): "Credit cycles," *Journal of political economy*, 105(2), 211–248.

Klenow, P. J., and H. Li (2022): "Entry Costs Rise with Growth," .

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.

Lanteri, A., and A. A. Rampini (2023a): "Constrained-efficient capital reallocation," *American Economic Review*, 113(2), 354–395.

——— (2023b): "Financing the Adoption of Clean Technology," .

Levine, R. (2005): "Chapter 12 finance and growth: Theory and evidence," *Handbook of economic growth*, 1(Part A), 865–934.

Lian, C., and Y. Ma (2021): "Anatomy of corporate borrowing constraints," *The Quar-*

terly 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.

MELITZ, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 71(6), 1695–1725.

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

MOLL, B. (2014): "Productivity losses from financial frictions: Can self-financing undo capital misallocation?," *American Economic Review*, 104(10), 3186–3221.

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

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.

QUERALTO, A. (2020): "A model of slow recoveries from financial crises," *Journal of Monetary Economics*, 114, 1–25.

RAMPINI, A. A. (2019): "Financing durable assets," *American Economic Review*, 109(2), 664–701.

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 Data Appendix

This appendix provides additional empirical results referenced in the main text.

A.1 Data Construction

Variables For the Compustat sample, 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 share*: ratio of research and development expense (`xrd`) to the sum of capital expenditures and research and development expense. *R&D-to-sales*: ratio of research and development expense to the average of sales (`sale`) in the previous 5 years.

3. *Patents*: Number of patents filed per year (based on the variable `filing_dated`) and market value of patents (based on the variable `xi_real`), constructed from the Kogan et al. (2017) dataset. To construct the patent-value-to-sales ratio, we use the average of sales (`sale`) in the previous 5 years.

4. *Net worth*: defined as sum of plant, property, and equipment and cash and short-term investments (`che`) minus total debt (sum of `dlc` and `dltt`).

5. *Cash flows*: measured as the sum of EBITDA and research and development expense divided by lagged plant, property, and equipment.

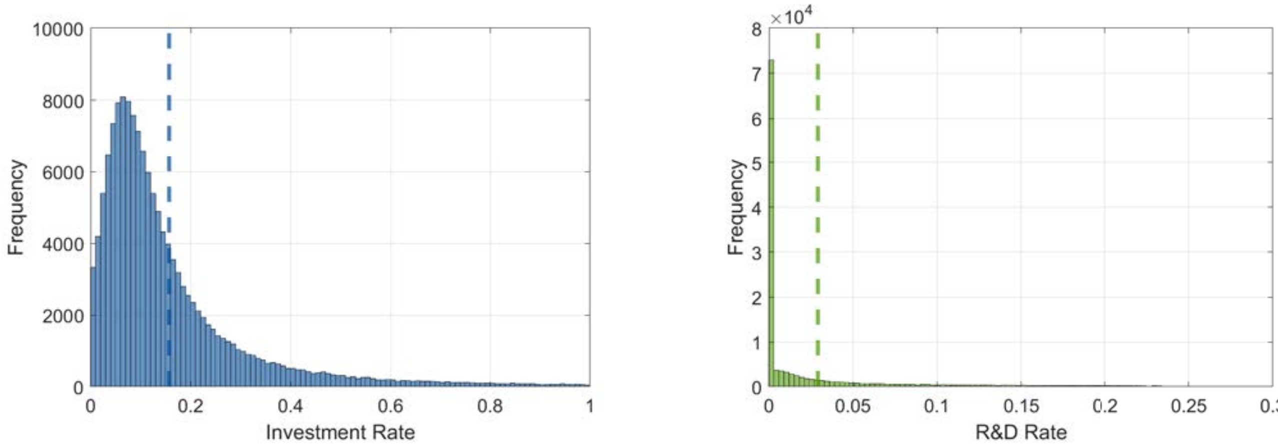
6. *Capital-to-employment*: defined at the ratio of lagged plant, property, and equipment to employment (`emp`).

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 ($\texttt{sic} = 9997$).

FIGURE A.1: 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 A.1.

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. R&D-to-sales ratios higher than 0.3.

   vi. Gross leverage (defined as the ratio of total debt to total assets) higher than 10 or negative.

Descriptive Statistics Table A.1 contains descriptive statistics of our final analysis sample. Figure A.1 plots the distribution of investment rates and R&D-to-sales ratios in our sample.

Table A.1
Descriptive statistics

	Mean	Median	St dev	95th	Observations
Investment rate	.155	.107	.15	.48	157,644
Investment spike	.233		.423		157,644
Investment rate | spike	.371	.311	.167	.778	36,735
Time since last spike	4.34	2	6.12	17	111,130
R&D share	.193	0	.29	.84	165,105
R&D-to-sales ratio	.027	0	.055	.166	133,054
Positive R&D expenditure	.433		.495		133,054
R&D-to-sales ratio | positive R&D expenditure	.063	.033	.069	.222	57,584
Leverage	.281	.243	.242	.736	166,641

Notes: This table shows descriptive statistics for variables used in the empirical analysis of Section 6.1. Investment rate, R&D-to-sales ratio, and leverage are defined in Appendix A.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 A.1.

A.2 Pecking Order for Additional Innovation Outcomes

In this appendix, we study the pecking order for a range of innovation outcomes, measures of size, and samples of firms. We have two main goals for this analysis. First, we compare to the existing literature, which tends to find that small firms do more innovation than large firms. While this finding seems inconsistent with our pecking order, we show that the differences result from focusing on different dimensions of the data, and that our pecking order holds for the dimensions most relevant to our model. Second, we show that our pecking order is robust to the majority of permutations that we consider, building confidence in our results.

Table A.2

Role of Within-Firm Variation For Various Measures of Size

	R&D share		R&D-to-sales		Patent activity		Patents per employee		Value of patents / sales	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	No FE	Firm FE	No FE	Firm FE	No FE	Firm FE	No FE	Firm FE	No FE	Firm FE
(a) Unrestricted sample										
log net worth	-0.021	0.021	-0.001	0.002	0.101	0.057	0.210	0.180	0.032	0.016
	(0.001)	(0.002)	(0.000)	(0.000)	(0.001)	(0.004)	(0.013)	(0.061)	(0.001)	(0.003)
log capital	-0.051	0.021	-0.006	0.001	0.089	0.066	-0.063	-0.063	0.026	0.016
	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.005)	(0.011)	(0.072)	(0.000)	(0.004)
log employment	-0.032	0.003	-0.005	0.002	0.102	0.084	-0.232	-0.392	0.024	0.008
	(0.001)	(0.003)	(0.000)	(0.001)	(0.001)	(0.005)	(0.012)	(0.076)	(0.000)	(0.003)
log sales	-0.032	0.014	-0.005	0.002	0.108	0.080	-0.061	-0.087	0.029	0.020
	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.005)	(0.012)	(0.077)	(0.000)	(0.004)
(b) Firms with more than 20 years of data										
log net worth	-0.011	0.024	0.001	0.002	0.103	0.049	0.256	0.272	0.039	0.021
	(0.001)	(0.003)	(0.000)	(0.001)	(0.002)	(0.007)	(0.018)	(0.096)	(0.001)	(0.005)
log capital	-0.019	0.026	-0.000	0.002	0.103	0.064	0.141	0.187	0.034	0.022
	(0.001)	(0.003)	(0.000)	(0.001)	(0.002)	(0.008)	(0.017)	(0.113)	(0.001)	(0.005)
log employment	-0.003	0.002	0.001	0.002	0.129	0.110	-0.011	-0.399	0.034	0.008
	(0.001)	(0.005)	(0.000)	(0.001)	(0.002)	(0.011)	(0.018)	(0.140)	(0.001)	(0.006)
log sales	-0.006	0.018	0.000	0.003	0.128	0.094	0.143	0.177	0.036	0.026
	(0.001)	(0.004)	(0.000)	(0.001)	(0.002)	(0.010)	(0.018)	(0.137)	(0.001)	(0.005)
(c) Continuously innovative firms with more than 20 years of data										
log net worth	-0.038	0.043	-0.000	0.003	0.117	0.050	0.227	0.494	0.052	0.039
	(0.001)	(0.005)	(0.000)	(0.001)	(0.003)	(0.011)	(0.027)	(0.179)	(0.001)	(0.009)
log capital	-0.051	0.050	-0.003	0.003	0.117	0.067	0.046	0.351	0.044	0.042
	(0.001)	(0.007)	(0.000)	(0.001)	(0.003)	(0.014)	(0.026)	(0.230)	(0.001)	(0.010)
log employment	-0.044	-0.005	-0.003	0.004	0.153	0.161	-0.255	-0.874	0.050	0.013
	(0.002)	(0.010)	(0.000)	(0.001)	(0.003)	(0.019)	(0.030)	(0.295)	(0.002)	(0.012)
log sales	-0.045	0.032	-0.003	0.005	0.138	0.130	-0.009	0.366	0.047	0.051
	(0.001)	(0.008)	(0.000)	(0.002)	(0.003)	(0.017)	(0.027)	(0.283)	(0.001)	(0.011)

Notes: Results from estimating the regression $o_{jt} = \alpha_j + \gamma \log s_{jt} + \varepsilon_{jt}$, where o_{jt} is the outcome of interest (R&D share, R&D-to-sales, indicator for positive patenting, patent per employee, or log market value per patent), s_{jt} is the measure of size (capital, sales, employment, or net worth), and α_j is a firm fixed effect. We standardize the size measures $\log s_{jt}$ over the entire sample. Standard errors, reported in parentheses, are clustered at the firm level. Panel (a) reports results using the sample of all firms and periods; panel (b) the sample of firms with at least 20 years of observations; and panel (c) the sample of firms with at least 20 years of observations and that are "continuously innovative" (i.e., firms that have conducted positive R&D or patenting activity over the last five years.)

Table A.2 contains all of this analysis. We estimate the regression (1) from the main text, reproduced here:

$$o_{jt} = \alpha_j + \gamma \log s_{jt} + \varepsilon_{jt}.$$

Relative to the main text, we expand the set of outcomes o_{jt} to include the R&D-to-sales ratio in columns (3) and (4) and the patents-per-employee ratio in columns (7) and (8). For each outcome, we report results with and without the firm fixed effects α_j in order to study the role of within-firm variation. We also perform this analysis for three samples of firms: the entire sample of firms in panel (a); firms with at least 20 years of data in panel (b), as in the main text; and Akcigit and Kerr (2018)'s sample of "continuously innovative firms," i.e., firms that have conducted positive R&D or patenting activity at some point over the last five years, in panel (c).

We draw three main conclusions from Table A.2. First, our use of within-firm variation variation is necessary to find that larger firms are more R&D-intensive. For either the R&D share or the R&D-to-sales ratios, the sign of the coefficient γ flips — i.e. large firms do less R&D, not more — when we exlude the firm fixed effects α_j. This result indicates that larger firms are less R&D-intensive than small firms on average, even though individual firms increase their R&D intensity as they grow. While the across-firm variation is an interesting source of firm heterogeneity, we do not focus on this dimension of the data because our model makes predictions about how financial frictions affect firms' decisions relative to what that same firm would have done absent those frictions. These predictions are most directly disciplined by variation in how the same firm has behaved at different points in time.[39]

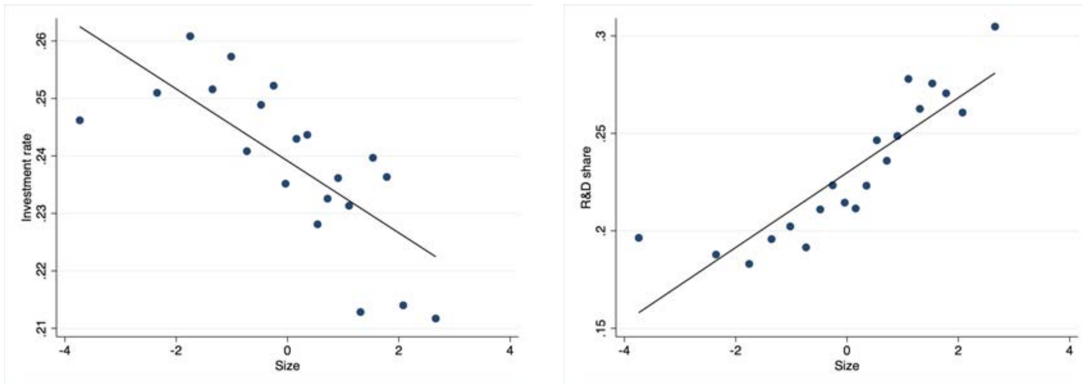
Second, Table A.2 shows that we are able to replicate Akcigit and Kerr (2018)'s finding that patents-per-employee increases in employment in our sample of Compustat firms. Akcigit and Kerr (2018) analyze a much wider set of firms from the Longitudinal Business Database (LBD) from the U.S. Census Bureau.[40] Therefore, the fact that we replicate their

[39]In contrast, the variation across firms also includes fixed differences across firms in things like production technologies. For example, large manufacturing firms may have a permanently high scale but engage in little innovation in order to focus on production, while small technology firms may have a permanently low optimal scale but focus on innovation rather than production. Consistent with this example, Appendix A.3 shows that our pecking order of firm growth also holds if we only use within-sector, rather than within-firm, variation.

[40]While the LBD contain information on sales and employment for a large set of firms, it does not contain information on capital investment. To our knowledge, the Census data that also contains investment only covers the manufacturing sector, as discussed in Footnote 6.

Figure A.2: The Pecking Order of Firm Growth in Orbis Data

Net investment rates R&D share



Notes: These figures report binned scatter plots of net investment rates and the share of firms with positive R&D by firms' size (measured by the log of real capital) in the Orbis-US dataset. All variables are demeaned at 4-digit-NAICS sector level. To construct the plots for investment rates and the share of firms with positive R&D rates, we add the unconditional mean of each variable to sector-level demeaned variables.

result in our smaller sample builds confidence about the external validity of our findings. However, this result does contrast with our finding that large firms are more innovation-intensive, even with the inclusion of fixed effects.

However, zooming out from only patents-to-employee as a function of employment, Table A.2 shows that our result that firms become more innovation-intensive as they grow holds for the majority of outcomes o_{jt} and measures of size s_{jt}. In fact, once we include firm fixed effects, the only combination of variables that goes in the opposite way is patents-to-employees as a function of employment, as discussed above. Looking across all outcomes, the pecking order is strongest when sorting firms by net worth, which is most closely linked to the shadow price of external finance in our model.

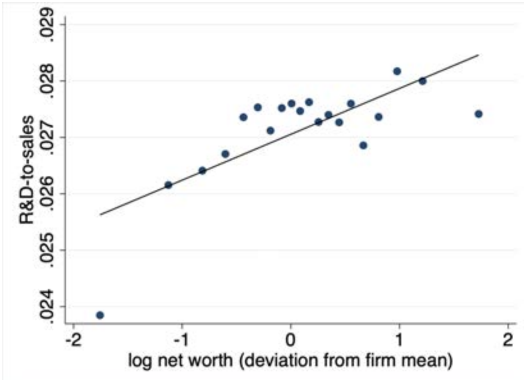
A.3 Robustness of Pecking Order

This section contains additional robustness analysis referenced in the main text.

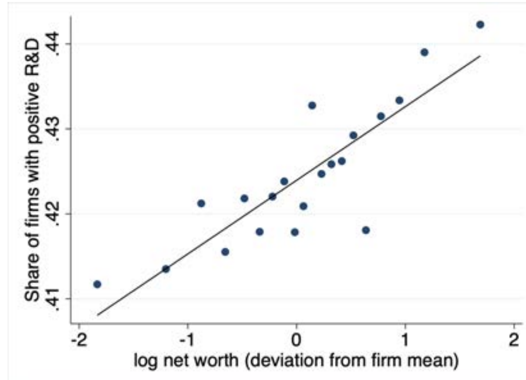
Orbis Data Figure A.2 shows our empirical results using Orbis, which includes data on privately held firms. In this analysis, we use within 4-digit sector variation rather than within-firm variation because the number of observations per firm is more limited for the sample of privately held firms (e.g., only 5% of privately held firms have at least 20 years of

Figure A.3: The Pecking Order of Firm Growth for Other Measures of Innovation

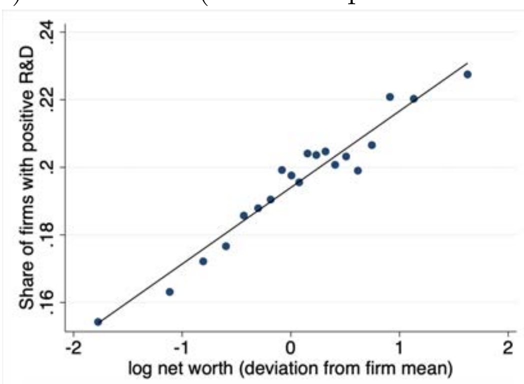
(a) R&D-to-sales



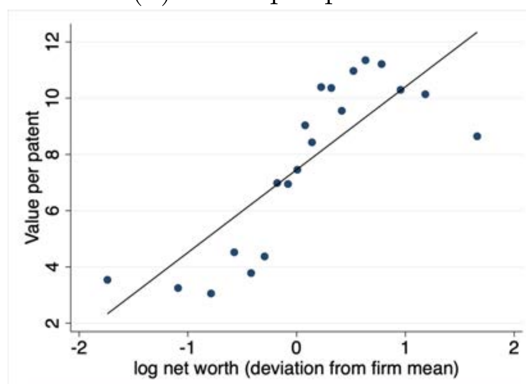
(b) R&D activity



(c) R&D share (after first positive R&D)



(d) Value per patent



Notes: Binned scatter plots of the R&D-to-sales ratio, the share of firms with positive R&D, the R&D share (conditional on already having an observation with positive R&D), and the market value per patent (computed following Kogan et al., 2017, and expressed in 1982 millions of dollars, deflated by the CPI) by the log of firm net worth. All variables are demeaned at the firm level. In order to make the units of the outcome variable more interpretable, we add back in the unconditional mean of the outcome variables across all firms. For variable definitions and sample selection, see Appendix A.

observations, which is the threshold we use to include firms in our baseline analysis). Because Orbis does not report gross investment expenditures, we instead infer net investment using the change in the value of the firm's capital stock. Similar to our results using Compustat data, we find that investment activity decreases as firms grow larger while innovation activity increases as firms grow larger.

Other Measures of Innovation Figure A.3 shows that our baseline bin-scatter plots look similar for four other measures of innovation inputs. Panel (a) plots the ratio of R&D expenditures to firm sales, which is often studied in the literature. Panel (b) plots an indicator variable for whether the firm reports positive R&D. Panel (c) plots the share

Table A.3
Sources of Variation in the Pecking Order

	(1) Investment rate	(2) R&D share	(3) R&D -to-sales	(4) R&D activity	(5) Patent activity	(6) Patent-value -to-sales
(a) Baseline						
$\hat{\gamma}$	-0.068	0.024	0.002	0.016	0.049	0.021
	(0.003)	(0.003)	(0.001)	(0.006)	(0.007)	(0.005)
N	45935	47286	41661	49105	49105	31176
R^2	0.263	0.857	0.877	0.854	0.639	0.678
(b) Sector fixed effects						
$\hat{\gamma}$	-0.014	0.019	0.006	0.073	0.155	0.062
	(0.002)	(0.005)	(0.002)	(0.010)	(0.010)	(0.012)
N	45939	47290	41664	49109	49109	31177
R^2	0.084	0.586	0.494	0.549	0.360	0.261
(c) Sector-by-time fixed effects						
$\hat{\gamma}$	-0.0005	0.014	0.007	0.085	0.196	0.071
	(0.003)	(0.005)	(0.002)	(0.011)	(0.011)	(0.014)
N	42962	44278	38785	46173	46173	28713
R^2	0.198	0.571	0.443	0.519	0.358	0.205
(d) Firm and time fixed effects						
$\hat{\gamma}$	-0.025	-0.007	0.003	0.018	0.097	0.021
	(0.003)	(0.004)	(0.001)	(0.008)	(0.009)	(0.005)
N	45935	47286	41661	49105	49105	31176
R^2	0.318	0.862	0.878	0.855	0.643	0.691

Notes: Panel (a) shows our baseline results, from estimating the regression $o_{jt} = \alpha_j + \gamma \log n_{jt} + \epsilon_{jt}$, where o_{jt} is the outcome of interest (investment rate, R&D share, R&D-to-sales, indicator for positive R&D, indicator for positive patenting, or patent-value-to-sales ratio); n_{jt} is net worth (standardized over the whole sample); and α_j is a firm fixed effect. Panel (b) reports the results from estimating $o_{jt} = \alpha_s + \gamma \log n_{jt} + \epsilon_{jt}$, where α_s is a sector fixed effect. Panel (c) reports the results from estimating $o_{jt} = \alpha_{st} + \gamma \log n_{jt} + \epsilon_{jt}$, where α_{st} is a sector-by-time fixed effect. Panel (d) reports the results from estimating $o_{jt} = \alpha_j + \alpha_t + \gamma \log n_{jt} + \epsilon_{jt}$, where α_t is a time fixed effect. Standard errors, reported in parentheses, are clustered at the firm level. For variable definitions and sample selection, see Appendix A.

of observations with positive R&D expenditures, conditional on already having reported positive R&D in the past. The hope is that this sample restricts to firms who have already set up the accounting infrastructure to report formal R&D and therefore has less measurement error. Finally, panel (d) plots the average market value per patent granted in a particular year, which is one measure of patent quality. In all cases, firms are more innovation-intensive when they have more net worth.

Sources of Variation Table A.3 shows that the pecking order is robust to using different sources of variation in the data. Panel (a) reports the regression coefficients (1) in our baseline specification from the main text. Panel (b) replaces the firm fixed effects α_j with 4-digit sector fixed effects to use within-sector, rather than within-firm, variation. Panel (c) then replaces those sector fixed effects with sector-by-year fixed effects to use within-sector and year variation. Finally, panel (d) uses both firm and year fixed effects (which absorb aggregate changes in the composition of investment and innovation). While the magnitudes of the coefficients change across specification, the basic patterns are robust to focusing on these alternative sources of variation.[41]

A.4 Innovation and Investment Spikes

This appendix contains four additional results about the relationship between R&D expenditures and investment spikes referenced in Section 6. First, Table A.4 Columns (1), (3), and (4) show that the main coefficient estimates are robust to including different numbers of lag H. Second, Table A.4 Column (2) shows that the result holds when spikes are defined as an investment rate of one standard deviation above the mean investment rate within sector s. Third, Table A.4 Column (5) shows that the results hold when adding size, sales growth, and current assets to the control vector X_{jt}. Finally, Figure A.4 complements the regression results with an event-study analysis around an investment spike.

B Model Appendix

This appendix provides various details of model analysis mentioned in the main text. Section B.1 characterizes firms' decision rules, proves Proposition 1, and provides details of the BGP. Section B.2 shows how to add labor to the model, as described in Footnote 11. Section B.3 provides details about the tax code discussed in Section 6.3 and Section 8.2. Section B.4

[41] The only exception is that the sign of the R&D share coefficient flips in panel (d), but this outlier estimate is only marginally significant.

Table A.4
Investment spikes and Innovation: Robustness

	(1)	(2)	(3)	(4)	(5)
$\frac{\text{RD}_{jt-1}}{\tilde{y}_{jt-1}}$	1.10	0.65	1.10	1.13	1.03
	(0.15)	(0.14)	(0.15)	(0.15)	(0.15)
$\frac{\text{cf}_{jt}}{k_{jt}}$	0.11	0.07	0.11	0.12	0.11
	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
years since spike$_{t-1}$	0.003	0.008	0.003	0.003	0.004
	(0.0008)	(0.0009)	(0.0008)	(0.0008)	(0.0008)
$\frac{k_{jt}}{n_{jt-1}}$	-0.05	-0.04	-0.05	-0.05	-0.05
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
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	55,647	39,215	55,647	50,117	54,191
Adj. R^2	0.300	0.220	0.300	0.294	0.314

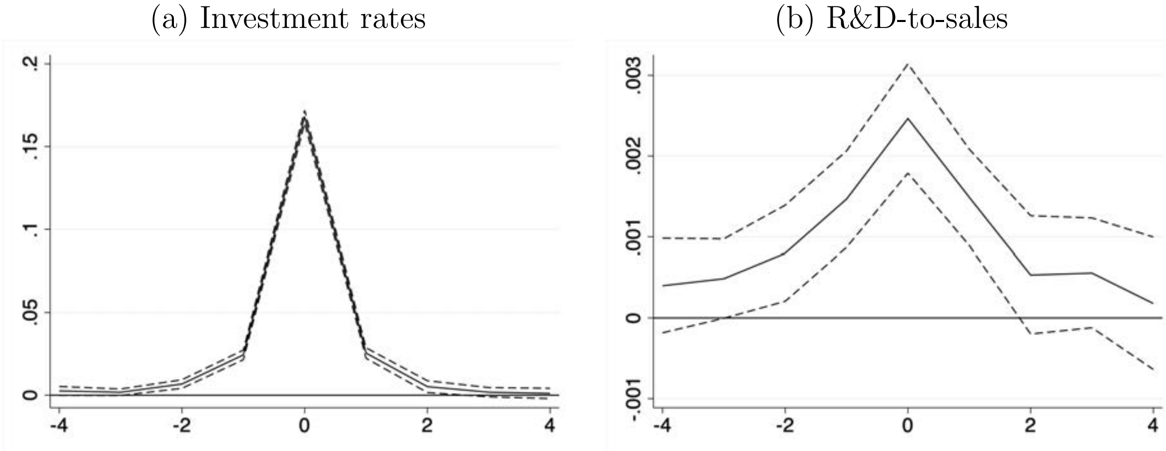
Notes: Results from estimating $\mathbb{1}\{\frac{x_{jt}}{k_{jt}} \geq \chi_s\} = \alpha_j + \alpha_{st} + \sum_{h=1}^{H} \beta_h \left(\frac{\text{RD}_{jt-h}}{\tilde{y}_{jt-h}}\right) + \Gamma' X_{jt} + \epsilon_{jt}$, where $\frac{x_{jt}}{k_{jt}}$ denotes the investment rate of firm j in period t; 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 residual. Column (1) reports estimates for the baseline specification of Table 2, 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 χ_{ts} 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 A.

derive the expressions relating to the costs of financial frictions in Section 7. Finally, Section B.5 sets up the planner's problem and proves Proposition 2.

B.1 Firms' Decision Rules and the BGP

This subsection 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 system, which is what we solve numerically. 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

FIGURE A.4: Event Study Analysis of Investment Spikes



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 $y_{jt} = \alpha_j + \alpha_{st} + \sum_{h=-4}^{4} \beta_h \mathbb{1}\{\frac{x_{jt+h}}{k_{jt+h}} \geq 0.2\} + \varepsilon_{jt}$, where y_{jt} denotes the investment rate ($\frac{x_{jt}}{k_{jt}}$) or R&D-to-sales ratio ($\frac{\text{RD}_t}{\tilde{y}_t}$); 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.

growth rate g in a balanced growth path.

B.1.1 Detrending

We will scale the problem by $Z_t = \left(A_t \int z_{jt} dj\right)^{\frac{1}{1-\alpha}} = \left(\int z_{jt} dj\right)^{\frac{1+a}{1-\alpha}}$. 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 (6) by Z_t to get

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

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

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 $\frac{k'}{Z_t} = \frac{k'}{Z_{t+1}} \frac{Z_{t+1}}{Z_t} = (1+g_t)\widetilde{k}'$ and $\frac{b'}{Z_t} = (1+g_t)\widetilde{b}'$. Now multiply and divide the continuation value by $\frac{Z_{t+1}}{Z_{t+1}}$ to get

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

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:

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

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 $\widetilde{n}' = \widetilde{z}'(\widetilde{k}')^\alpha + (1-\delta)\widetilde{k}' - \widetilde{b}'$. 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.

B.1.2 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, for convenience, we un-detrend those first-order conditions to get the system of equations in the second part of Proposition 1.

Lagrangian The Lagrangian of the detrended Bellman equation (16) is

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

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 $MPK(\widetilde{z}',\widetilde{k}') = \alpha\widetilde{z}'(\widetilde{k}')^{\alpha-1}$ is the marginal product of capital. Using very similar steps to above, the terms in the continuation value can be collected to yield

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

The first-order condition for innovation i is

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

Consider the term in the continuation value in the case where the firm exits in the next period. We can write this expectation as $\mathbb{E}_t[\widetilde{n}'] = \eta(i)\mathbb{E}^{\varepsilon}[\widetilde{n}'|\iota = 1] + (1-\eta(i))\mathbb{E}^{\varepsilon}[\widetilde{n}'|\iota = 0]$ where $\mathbb{E}^{\varepsilon}$ denotes the expectation over the idiosyncratic shocks ε. Hence, we have $\frac{\partial \mathbb{E}_t[\widetilde{n}']}{\partial i} = \eta'(i)\left(E^{\varepsilon}[\widetilde{n}'|\iota = 1] - E^{\varepsilon}[\widetilde{n}'|\iota = 0]\right)$. By a similar argument,

$$\frac{\partial \mathbb{E}_t[\widetilde{v}_{t+1}(\widetilde{z}', \widetilde{n}')]}{\partial i} = \eta'(i)\left(E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}', \widetilde{n}')|\iota = 1] - E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}', \widetilde{n}')|\iota = 0]\right).$$

Putting these all together yields

$$(1+\lambda_t(\widetilde{z}, \widetilde{n}))\widetilde{z}^{\frac{1}{1-\alpha}} \geq \frac{1+g_t}{1+r_t}\eta'(i)\begin{bmatrix} \pi_d\left(E^{\varepsilon}[\widetilde{n}'|\iota = 1] - E^{\varepsilon}[\widetilde{n}'|\iota = 0]\right) + \\ (1-\pi_d)\left(E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}', \widetilde{n}')|\iota = 1] - E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}', \widetilde{n}')|\iota = 0]\right)\end{bmatrix}, \quad (20)$$

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:

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

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

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

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:

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

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:

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

where $\widetilde{n}' = \widetilde{z}'(\widetilde{k}_t'^*(\widetilde{z}))^\alpha + (1-\delta)\widetilde{k}_t'^*(\widetilde{z}) - \widetilde{b}_t'^*(\widetilde{z})$ 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

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

$$1 \geq \frac{\eta'(i)}{\widetilde{z}^{\frac{1}{1-\alpha}}}\frac{1+g_t}{1+r_t}\mathbb{E}_t\begin{bmatrix}\pi_d\left(E^\varepsilon[\widetilde{n}'|\iota=1] - E^\varepsilon[\widetilde{n}'|\iota=0]\right) + \\ (1-\pi_d)\left(E^\varepsilon[\widetilde{v}_{t+1}^*(\widetilde{z}')|\iota=1] - E^\varepsilon[\widetilde{v}_{t+1}^*(\widetilde{z}')|\iota=0]\right)\end{bmatrix}. \quad (24)$$

Note that the innovation policy implicitly enters the first-order condition for capital (23) through the expectations operator. Nevertheless, one can verify from (23) and (24) 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:

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

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

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

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

Un-Detrending the Conditions We now show that the detrended first-order conditions (18), (19), and (20) derived above imply the conditions (7), (8), and (9) 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

$$MPK_{t+1}(z',k') = \alpha\frac{A_{t+1}z'}{(k')^{1-\alpha}} = \alpha\frac{A_{t+1}z'/Z_{t+1}^{1-\alpha}}{(k')^{1-\alpha}/Z_{t+1}^{1-\alpha}} = \alpha\widetilde{z}(\widetilde{k}')^{1-\alpha}.$$

Hence, the detrended first-order condition (19) directly implies the undetrended first-order condition (8) (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

$$(1+\lambda_t(z,n))Z_t\widetilde{z}^{\frac{1}{1-\alpha}} \geq \frac{\eta'(i_t(z,n))}{1+r_t}Z_{t+1}\mathbb{E}_t\begin{bmatrix}\pi_d\left(E^{\varepsilon}[\widetilde{n}'|\iota=1] - E^{\varepsilon}[\widetilde{n}'|\iota=0]\right) + \\ (1-\pi_d)\left(E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\iota=1] - E^{\varepsilon}[\widetilde{v}_{t+1}(\widetilde{z}',\widetilde{n}')|\iota=0]\right)\end{bmatrix}$$

By definition of the detrended variables, this equation is the same as the un-detrended condition (9) from the main text. The nonnegativity constraint for dividends (7) follows directly from our detrending of the problem.

B.1.3 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 $\widetilde{n}' = \widetilde{z}'\widetilde{k}_t'(\widetilde{z},\widetilde{n})^{\alpha} + (1-\delta)\widetilde{k}_t'(\widetilde{z},\widetilde{n}) - \widetilde{b}_t'(\widetilde{z},\widetilde{n})$ is the law of motion for detrended state variables induced by the policy rules.[42]

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

[42]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.

shown that the firm value function and decision rules are all scaled by Z_t in the sense that 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:

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

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

B.2 Adding Labor to the Model

Adding labor extends the model in two ways. First, as discussed in the main text, the production function becomes $y_{jt} = A_t z_{jt} k_{jt}^{\alpha} \ell_{jt}^{\nu}$, 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.[43]

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:

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

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

[43]Given these additively separable preferences over consumption and labor supply, balanced growth requires log utility over consumption. We could alternatively allow for a non-unitary EIS if we instead assume preferences fall within the more general King, Plosser and Rebelo (1988) class.

labor demand and simplifying yields

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

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:

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

B.3 Incorporating Corporate Taxes and Bonus Depreciation

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.[44] 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

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

well. Therefore, their total tax bill in a given period is

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

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 value of tax deductions per unit of investment:*

$$\widehat{\zeta}_t = \sum_{\iota=0}^{\infty} \left(\prod_{p=0}^{s} \frac{1}{1+r_{t+p}} \right) (1-\widehat{\delta})^s. \tag{27}$$

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

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

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

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 $\widehat{\zeta}_t$. Temporary tax incentives, like the Bonus Depreciation Allowance, temporarily increase $\widehat{\zeta}_t$ and therefore act as shocks to the relative price of investment. The TCJA 2017 tax reform introduced full expensing, which increased $\widehat{\zeta}_t = 1$ because it allows firms to fully deduct investment expenditures from their tax bill in the period they are incurred. To keep our analysis simple, we will directly work with the composite shock $\zeta_t = \tau\widehat{\zeta}_t$ and assume that $\zeta_t = 0$ in the balanced growth path. This assumption implies that we do not have to recalibrate the model to accommodate steady state taxes. Instead, we will calibrate our tax shocks ζ_t as deviations from their initial value.

B.4 Aggregate Costs of Financial Frictions

We derive the two results referenced in Section 7 of the main text.

Approximation of the Long-Run Growth Rate The long-run growth rate is given by $1 + g = (1 + \widetilde{g})^{\frac{1+a}{1-\alpha}}$, where $\widetilde{g}$ is the growth rate of average productivity $\int z_{jt}dj$. In Appendix C, we show that

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

where $\Phi_t(s)$ is the distribution of firms over individual states $s = (z, n)$ in period t along a balanced growth path, and $z'(s, \varepsilon') = ze^{\varepsilon'}e^{\Delta}$ with probability $\eta(i(s))$ and $z'(s, \varepsilon') = ze^{\varepsilon'}$

with probability $1-\eta(i(s))$. Plug this in to get

$$\begin{aligned}
1+\widetilde{g} &= \frac{\int \left(\eta(i(s))ze^{\varepsilon'}e^{\Delta}+(1-\eta(i(s)))ze^{\varepsilon'}\right)p(\varepsilon')\Phi_t(s)d\varepsilon' ds}{\int z\Phi_t(s)ds} \\
\implies 1+\widetilde{g} &= \frac{\int e^{\varepsilon'}p(\varepsilon')d\varepsilon' \times \int \left(\eta(i(s))ze^{\Delta}+(1-\eta(i(s)))z\right)\Phi_t(s)ds}{\int z\Phi_t(s)ds} \\
\implies 1+\widetilde{g} &= \frac{e^{\sigma_\varepsilon^2/2}\times \int \left(1+\eta(i(s))(e^{\Delta}-1)\right)z\Phi_t(s)ds}{\int z\Phi_t(s)ds} \\
\implies 1+\widetilde{g} &= e^{\sigma_\varepsilon^2/2}\left(\frac{\int z\Phi_t(s)ds}{\int z\Phi_t(s)ds}+(e^{\Delta}-1)\frac{\int z\eta(i(s))\Phi_t(s)ds}{\int z\Phi_t(s)ds}\right) \\
\implies 1+\widetilde{g} &\approx \left(1+(e^{\Delta}-1)(\int \eta(i(s))\Phi_t(s)ds)\frac{\int z\Phi_t(s)ds}{\int z\Phi_t(s)ds}\right) \\
\implies \widetilde{g} &\approx (e^{\Delta}-1)\int \eta(i(s))\Phi_t(s).
\end{aligned}$$

Hence, we have

$$\begin{aligned}
1+g &= (1+\widetilde{g})^{\frac{1+a}{1-\alpha}} \\
\implies \log(1+g) &= \frac{1+a}{1-\alpha}\log(1+\widetilde{g}) \\
\implies g &\approx \frac{1+a}{1-\alpha}(e^{\Delta}-1)\int \eta(i(s))\Phi_t(s),
\end{aligned}$$

as in the main text.

Upper Bound on Misallocation Costs The upper bound on misallocation costs in the main text compares actual output along the BGP, Y_t, to the benchmark

$$Y_t^* = \max_{k_{jt}} A_t \int z_{jt}k_{jt}^{\alpha}dj \text{ such that } \int k_{jt}dj \le K_t.$$

The first-order condition with respect to k_{jt} can be rearranged to

$$k_{jt} = \left(\frac{\alpha A_t z_{jt}}{\lambda}\right)^{\frac{1}{1-\alpha}}, \tag{28}$$

where, abusing notation, λ is the Lagrange multiplier on the constraint $\int k_{jt}dj \le K_t$.

Integrating (28) across firms j and using $\int k_{jt}dj = K_t$ gives

$$\lambda = \frac{\alpha A_t}{K_t^{1-\alpha}} \left(\int z_{jt}^{\frac{1}{1-\alpha}} dj \right)^{1-\alpha}.$$

Plug this expression into the FOC (28) and rearrange to get

$$\frac{k_{jt}}{K_t} = \frac{z_{jt}^{\frac{1}{1-\alpha}}}{\int z_{jt}^{\frac{1}{1-\alpha}} dj}.$$

Aggregate TFP in this allocation is therefore

$$TFP_t^* = A_t \int z_{jt} \left(\frac{k_{jt}}{K_t} \right)^{\alpha} dj$$
$$\implies TFP_t^* = A_t \int z_{jt} \left(\frac{z_{jt}^{\frac{1}{1-\alpha}}}{\int z_{jt}^{\frac{1}{1-\alpha}} dj} \right)^{\alpha}$$
$$\implies TFP_t^* = A_t \frac{\int z_{jt}^{\frac{1}{1-\alpha}}}{\left(\int z_{jt}^{\frac{1}{1-\alpha}} \right)^{\alpha} dj}$$
$$\implies TFP_t^* = A_t \left(\int z_{jt}^{\frac{1}{1-\alpha}} dj \right)^{1-\alpha}.$$

Taking the ratio of this to actual TFP gives the expression in the main text.

B.5 Planner's Problem and Proof of Proposition 2

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

$$W_t(\Phi) = \max_{k'(\cdot),i(\cdot),b'(\cdot)} \frac{C^{1-\sigma}-1}{1-\sigma} + \beta W_{t+1}(T(\Phi;k'(\cdot),i(\cdot),b'(\cdot))) \text{ such that} \tag{29}$$

$$C = \int [Azk^{\alpha} + (1-\delta)k]\,\Phi(s)ds - (1-\pi_d)\int \left[k'(s) + (Az)^{\frac{1}{1-\alpha}} i(s)\right]\Phi(s)ds \tag{30}$$
$$- \pi_d \int k'\Phi^0(z',k',b')dz'dk'db'$$

$$Azk^{\alpha} + (1-\delta)k - b - k'(s) - (Az)^{\frac{1}{1-\alpha}} i(s) + \frac{b'(s)}{1+r_t} \geq 0 \text{ for all } s \tag{31}$$

$$b'(s) \leq \theta k'(s) \text{ for all } s \tag{32}$$

$$A = \left(\int z\Phi(s)dz\right)^{a} \tag{33}$$

$$T(\Phi;k'(\cdot),i(\cdot),b'(\cdot))(z',k',b') = \pi_d \Phi^0(z',k',b') \tag{34}$$
$$+ (1-\pi_d)\int \begin{bmatrix} (\mathbb{1}\{k'=k'(s)\} \times \mathbb{1}\{b'=b'(s)\})\times \\ (\eta(i(s))\mathbb{1}\{z'=ze^{\Delta}e^{\varepsilon}\} + (1-\eta(i(s)))\mathbb{1}\{z'=ze^{\varepsilon}\}) \end{bmatrix} p(\varepsilon)\Phi(s)ds,$$

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 (29) 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, $\frac{\delta W}{\delta \Phi(\bar{s})}(\Phi)$ denotes the Gâteaux derivative with respect to the mass of households at point s, which itself is a function of the entire distribution Φ.[45] The time subscripts reflect the dependence on the path of the real interest rate in firms' borrowing decisions. For notational simplicity we will often omit the dependence on Φ and the time subscripts.

We will use these tools to solve the planner's problem (29) using Lagrangian methods. Let $\lambda(s)$ denote the multiplier on the no-equity issuance constraint (31), $\mu(s)$ denote the

[45] 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.

multiplier on the collateral constraint (32), and Λ denote the multiplier on the innovation externality (33). We will directly plug in the definitions of consumption (30) and the transition function for the distribution (34). With all this notation in hand, the Lagrangian is

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

where it is understood that C and $T(\Phi; k'(\cdot), i(\cdot), b'(\cdot))$ stand in for (30) and (34).

We proceed in two steps. First, subsection B.5.1 takes the first-order conditions with respect to all the planner's choices. Second, subsection B.5.2 characterizes those choices in terms of the marginal social value function from Proposition 2 in the main text.

B.5.1 First Order Conditions

We analyze each first-order condition separately.

Aggregate productivity The FOC with respect to aggregate productivity is

$$C^{-\sigma}\left[\int zk^{\alpha}\Phi(s)ds - \frac{1-\pi_d}{1-\alpha}\int A^{\frac{\alpha}{1-\alpha}} z^{\frac{1}{1-\alpha}} i(s)\Phi(s)ds\right]$$
$$+ \int \lambda(s)\left[zk^{\alpha}\Phi(s)ds - \frac{1}{1-\alpha}A^{\frac{\alpha}{1-\alpha}} z^{\frac{1}{1-\alpha}} i(s)\right] ds = \Lambda. \tag{35}$$

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 $\widetilde{\Lambda} = \frac{\Lambda}{C^{-\sigma}}$.[46] Plugging these in and simplifying yields

$$\widetilde{\Lambda} = \pi_d \int zk^{\alpha}\Phi(s)ds + (1-\pi_d)\int (1+\widetilde{\lambda}(s))\left[zk^{\alpha}\Phi(s)ds - \frac{1}{1-\alpha}A^{\frac{\alpha}{1-\alpha}} z^{\frac{1}{1-\alpha}} i(s)\right]\Phi(s)ds. \tag{36}$$

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

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

[46] 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 (34). 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

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

Plugging this into the FOC gives

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

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

$$(Az)^{\frac{1}{1-\alpha}}(1+\widetilde{\lambda}(s)) = \frac{\beta}{C^{-\sigma}}\eta'(i(s)) \left[\int \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\Delta}e^{\varepsilon}, k'(s), b'(s))} p(\varepsilon) d\varepsilon - \int \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\varepsilon}, k'(s), b'(s))} p(\varepsilon) d\varepsilon \right]. \quad (37)$$

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 and noting that only the part of the transition function from incumbents will matter for the derivatives 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

$$1 + \widetilde{\lambda}(s) = \theta \widetilde{\mu}(s) + \frac{\beta}{C^{-\sigma}} \begin{bmatrix} \eta(i(s)) \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\Delta} e^{\varepsilon}, k'(s), b'(s)))} p(\varepsilon) d\varepsilon + \\ (1-\eta(i(s))) \int \frac{\delta}{\delta k'(s)} \frac{\delta W(\Phi')}{\delta \Phi'(ze^{\varepsilon}, k'(s), b'(s)))} p(\varepsilon) d\varepsilon \end{bmatrix} \tag{38}$$

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

$$\frac{\widetilde{\lambda}(s)}{1+r_t}=\widetilde{\mu}(s)-\frac{\beta}{C^{-\sigma}}\left[\begin{array}{l}\eta(i(s))\int \frac{\delta}{\delta b'(s)}\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\Delta}e^{\varepsilon},k'(s),b'(s)))}p(\varepsilon)d\varepsilon+\\(1-\eta(i(s)))\int \frac{\delta}{\delta b'(s)}\frac{\delta W(\Phi')}{\delta\Phi'(ze^{\varepsilon},k'(s),b'(s)))}p(\varepsilon)d\varepsilon\end{array}\right] \tag{39}$$

B.5.2 Marginal Social Value Functions

The optimal choices to the planner's problem are given the FOCs (36), (37), (38), and (39), together with the complementarity slackness conditions. In order to arrive at the results in Proposition 2, 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

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

From the definition of the transition function (34), 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

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

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

$$\begin{aligned}\widetilde{\omega}(s;\Phi) =& \pi_d \left[Azk^{\alpha} + (1-\delta)k + \widetilde{\Lambda} a \left(\int z\Phi(s)ds \right)^{a-1} z \right] + \\ &+ (1-\pi_d) \begin{bmatrix} Azk^{\alpha} + (1-\delta)k - k'(s) - (Az)^{\frac{1}{1-\alpha}} i(s) + \widetilde{\Lambda} a \left(\int z\Phi(s)ds \right)^{a-1} z \\ + \beta \left(\frac{C'}{C} \right)^{-\sigma} \mathbb{E}^{\varepsilon} \left[\eta(i(s))\widetilde{\omega}(s';\Phi') + (1-\eta(i(s)))\widetilde{\omega}(s';\Phi') \right] \end{bmatrix}.\end{aligned} \tag{40}$$

We are finally in a position to derive the equations in Proposition 2 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

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

be the planner's social marginal value function plus the path of borrowing and debt repayments starting from period t. Plugging this into (40) gives the *augmented Bellman equation*

$$\begin{aligned}\widehat{\omega}_t(s) =& \pi_d \left[Azk^{\alpha} + (1-\delta)k - b + \widetilde{\Lambda} a \left(\int z\Phi(s)ds \right)^{a-1} z \right] + \\ &+ (1-\pi_d) \begin{bmatrix} Azk^{\alpha} + (1-\delta)k - b - k'(s) - (Az)^{\frac{1}{1-\alpha}} i(s) + \widetilde{\Lambda} a \left(\int z\Phi(s)ds \right)^{a-1} z \\ + \frac{b'(s)}{1+r_t} + \beta \left(\frac{C'}{C} \right)^{-\sigma} \mathbb{E}^{\varepsilon} \left[\eta(i(s))\widehat{\omega}_t(s') + (1-\eta(i(s)))\widehat{\omega}_t(s') \right] \end{bmatrix}.\end{aligned} \tag{41}$$

To keep notation even simpler, define $\widehat{\Lambda} = \widetilde{\Lambda} a \left(\int z\Phi(s)ds \right)^{a-1}$ and 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 (41) and let $\widehat{\omega}_t^{\text{cont}}$ second set of brackets in (41). 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

$$\widehat{\omega}_t^{\text{cont}}(s) = Azk^{\alpha} + (1-\delta)k - b - k'(s) - (Az)^{\frac{1}{1-\alpha}} i(s) + \frac{b'(s)}{1+r_t} + \widehat{\Lambda} z + \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \mathbb{E}_t \left[\widehat{\omega}_{t+1}(s') \right] \tag{42}$$

This Bellman-like equation (42) is similar to the augmented Bellman equation (13) from Proposition 2 except that (42) 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 (42) 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.

C Numerical 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. Given a guess of g, we solve for individual firms' decision rules, computed the implied growth from those decision rules, and check whether that implied growth is consistent with our guess for g. For each candidate growth rate, the most difficult part is solving for the individual decisions.

Individual decisions Given a guess for the growth rate g, 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 $\widetilde{k}'^{*}_{(\mathrm{it})}(\widetilde{z})$, $i^{*}_{(\mathrm{it})}(\widetilde{z})$, and $\widetilde{v}_{(\mathrm{it})}(\widetilde{z})$, 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 $\widetilde{k}'^{*}_{(\mathrm{it})+1}(\widetilde{z}) = \left(\alpha \frac{\mathbb{E}_t[\widetilde{z}']}{r-\delta}\right)^{\frac{1}{1-\alpha}}$, 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 $i^{*}_{(\mathrm{it})}(\widetilde{z})$ to evaluate the expectation.

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

$$i^*_{(\text{it})+1}(\widetilde{z}) = \max\{0, \eta'^{-1}\left(\widetilde{z}^{\frac{1}{1-\alpha}}\frac{1+r}{1+g}\begin{bmatrix}\pi_d\left(E^{\varepsilon}[\widetilde{n}'|\iota=1]-E^{\varepsilon}[\widetilde{n}'|\iota=0]\right)+ \\ (1-\pi_d)\left(E^{\varepsilon}[\widetilde{v}^*_{(\text{it})}(\widetilde{z}')|\iota=1]-E^{\varepsilon}[\widetilde{v}^*_{(\text{it})}(\widetilde{z}')|\iota=0]\right)\end{bmatrix}^{-1}\right)\}.$$

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. We pre-compute the inverse function $\eta'^{-1}(y)$.

(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 the 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:

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

  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:

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

(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 $\widetilde{b}'=\theta\widetilde{k}'$:

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

- 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:

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

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).

Updating guess of the growth rate g We now need to compute the aggregate growth rate implied by these decision rules. We compute the growth rate of average productivity $\int z_{jt}dj$, $\widetilde{g}$, using the definition

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

where $s=(z,n)$ denotes the individual state vector. 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

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

The numerator in this integral is

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

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

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

Given this value of $\widetilde{g}$, we can then compute the implied growth of Z_t as $1+\widehat{g} = (1+\widetilde{g})^{\frac{1+a}{1-\alpha}}$.

Taken together, this procedure defines a mapping from the current guess of the growth rate, g, to a new guess, $\widehat{g} = f(g)$. The balanced growth path is a fixed point of this mapping. We compute the fixed point using a nonlinear equation solver to numerically solve the equation $\widehat{g} - f(g) = 0$.

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 $h(\{g_t, r_t\}) = 0$, where h 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 $h(\{g_t, r_t\})$ are then the aggregate consistency conditions:

$$\frac{1}{\beta}\left(\frac{C_{t+1}}{C_t}\right)^{\sigma} - (1 + r_t) = 0.$$

D Additional Quantitative Results

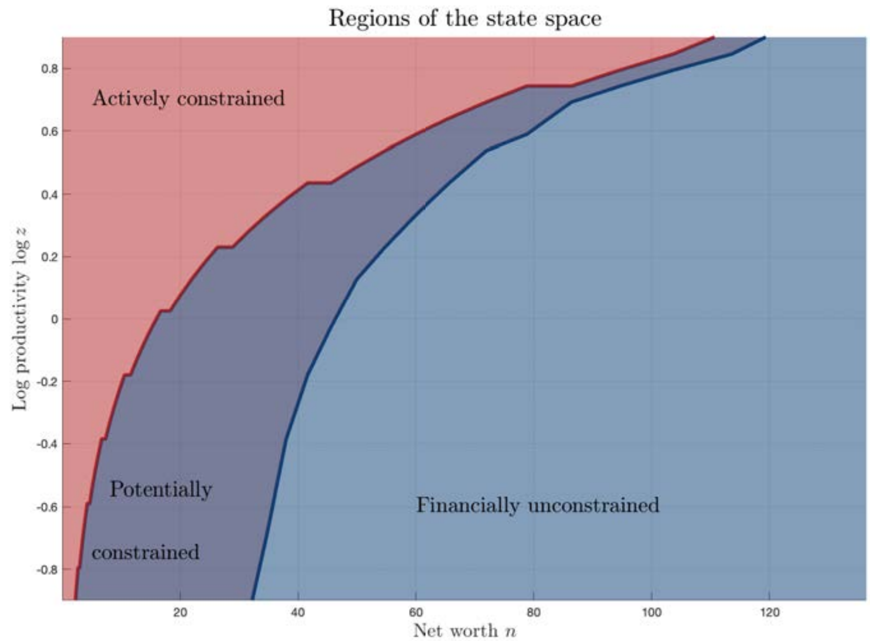
This appendix provides details of quantitative results described in the main text.

D.1 Sources of Firm Heterogeneity

Figure E.1 visualizes the partition of the state space characterized in Proposition 1 in the detrended BGP. 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 constrained.

Decision Rules Figure E.2 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 firm growth from Section 5, 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 innovation. While constrained,

FIGURE E.1: Partition of the State Space



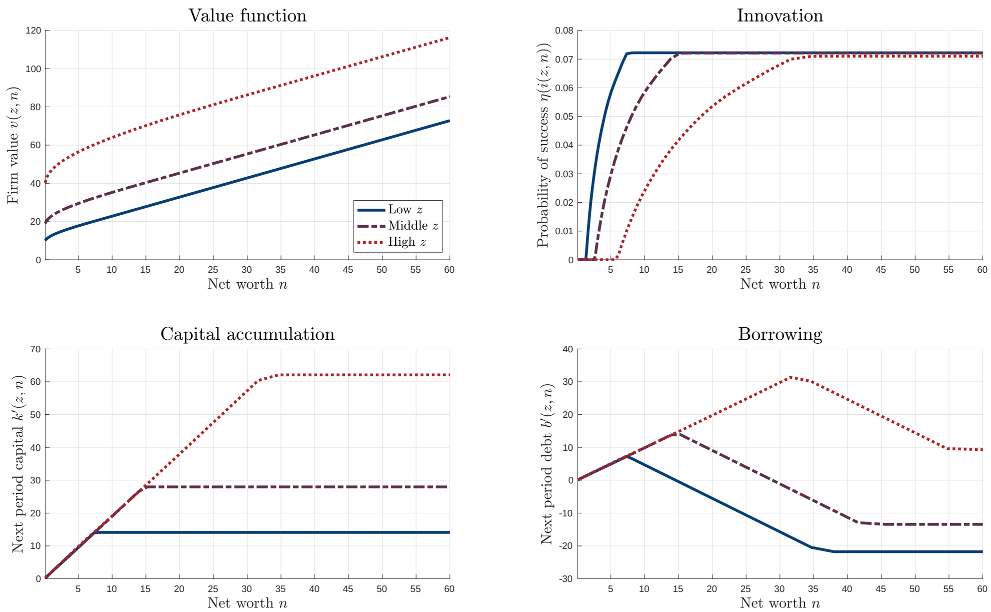
Notes: partition of the state space from Proposition 1 in the market BGP. Net worth n and log productivity $\log z$ have been detrended following Appendix B.

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 of how the cost of innovation is scaled by productivity.

Figure E.3 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 E.2). Among constrained firms, 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 $b' = \theta k'$

$$k'(z,n) = n + \frac{\theta k'(z,n)}{1+r} \implies \frac{\partial k'(z,n)}{\partial n} = \left(1 - \frac{\theta}{1+r}\right)^{-1} \approx 2,$$

FIGURE E.2: Decision Rules



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

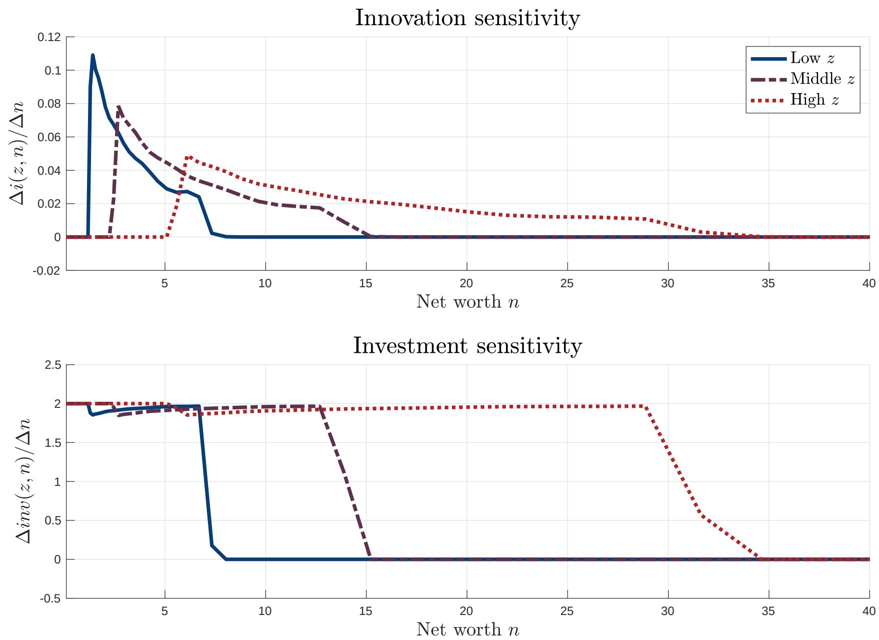
where the last approximation uses our calibrated values of $\theta = 0.52$ and $r = 0.04$. Since firms can lever up investment with borrowing, their investment-cash flow sensitivities are above one. Constrained firms with positive innovation have a smaller investment-cash flow sensitivity because they put some of the additional funds toward innovation as well:

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

Quantitatively, Figure E.3 shows that the innovation-cash flow sensitivities are an order of magnitude smaller than the investment-cash flow sensitivities.

Lifecycle Dynamics Figure E.4 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 years of life, the firm has a very high investment rate and does not innovate. 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 1 in the main

Figure E.3: 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.

text. In this particular sample path, the firm receives two successful innovations, both of which successful innovations are accompanied by investment spikes.

Decomposing the Return to Capital Figure E.5 decomposes the return to capital from the pecking order plot Figure 3 as well as its "MPK component"

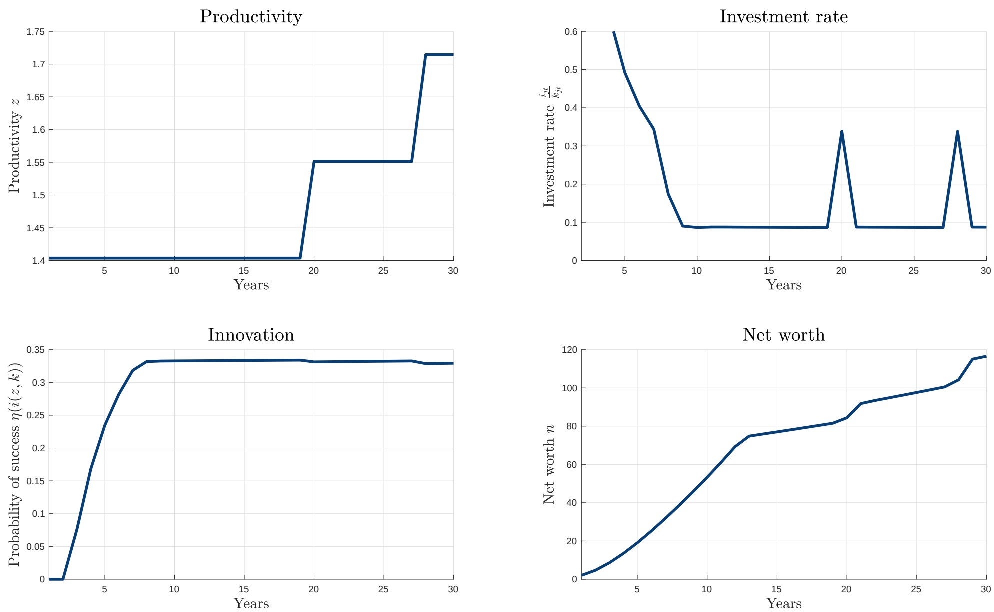
$$\frac{1}{1+r_t}\mathbb{E}_t\left[(MPK_{t+1}(z',k')+1-\delta)\times(1+(1-\pi_d)\lambda_{t+1}(z',n'))\right],$$

i.e. the part of the return to capital not driven by its value as collateral $\theta\mu_t(z,n)$. The figure shows that the majority of the difference between the return to capital vs. innovation is due to the MPK component, not its collateral value.

D.2 Distribution of Investment, Innovation, and Leverage

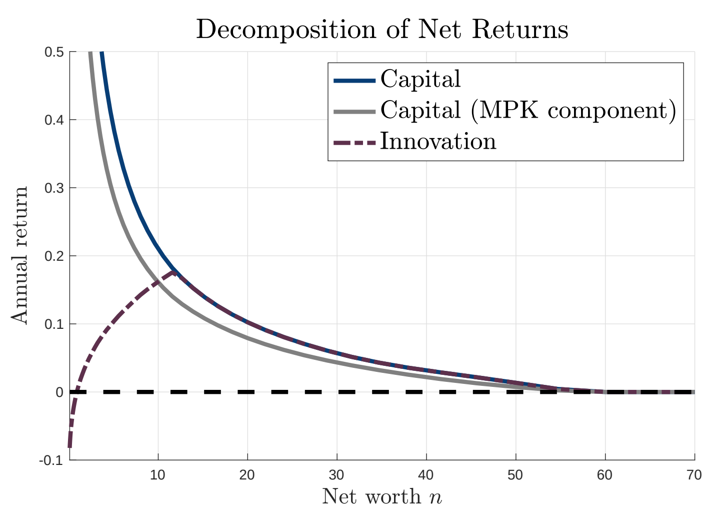
Table E.1 compares a number of moments of the stationary distribution of investment, R&D, and leverage from our model to their counterparts in the Compustat data. The model

Figure E.4: 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 .

Figure E.5: Decomposing the Return to Capital



Notes: the return to investment and innovation, defined as the RHS of Euler equations (8) and (9) minus 1. “Capital (MPK component)” refers to the return to capital excluding the collateral value $\theta\mu_t(z, n)$.

Table E.1
Distribution of Investment, Innovation, and Leverage

Statistic	**Data**	**Model**
Investment spending		
$\mathbb{E}[x_{jt}/k_{jt}]$	0.15	0.16
$\sigma(x_{jt}/k_{jt})$ (targeted)	0.15	0.13
$\mathbb{E}[x_{jt}/k_{jt}\|\text{spike}]$ (targeted)	0.37	0.35
R&D spending		
$\mathbb{E}[\text{RD}_{jt}/y_{jt}]$	0.06	0.03
$\text{Frac}(\text{RD}_{jt}/y_{jt} > 0)$	0.45	0.92
$\mathbb{E}[\text{RD}_{jt}/y_{jt}\|\text{RD}_{jt} > 0]$ (targeted)	0.06	0.06
Leverage		
Mean gross leverage, all (targeted)	0.34	0.30
Mean gross leverage, Compustat	0.21	0.28
SD gross leverage, Compustat	0.22	0.24
Mean net leverage, Compustat	0.13	0.09
SD net leverage, Compustat	0.33	0.36

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 $\max\{b_{jt}, 0\}$.

endogenously matches the average investment rate fairly well even though they it was not directly targeted in the calibration. The model model also matches the first two moments of leverage fairly well, either in terms of gross or net leverage. However, the model over-predicts the share of firms with positive R&D spending compared to the data. We choose not to target this statistic because it is well-known that firms under-report R&D expenditures, especially along the extensive margin.

D.3 Investment Tax Shocks

The Bonus Depreciation Allowance allowed 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 $\Delta\zeta_t = b_t(1 - \zeta)$ where $\zeta < 1$ is the present value of deductions under the baseline schedule.

Table E.2

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
$\frac{1-\tau\zeta_{st}}{1-\tau}$	-1.37	-1.73	-0.14	-0.20
	(0.16)		(0.04)	
R^2	0.35		0.88	

Notes: estimates of $\widehat{\gamma}$ from the regression (43) in columns (1) and (2) or from the regression (44) in columns (3) - (6). Standard errors, reported in parentheses, are clustered by firms. "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 0.8 (giving a half-life of roughly two years).

Zwick and Mahon (2017) show that sectoral heterogeneity in the baseline tax depreciation schedule across sectors, ζ_s, provides exogenous variation that can be used to identify the effect of the Bonus, $\Delta\zeta_{st} = b_t(1-\zeta_s)$, on investment. We estimate their specification in our Compustat sample with the regression

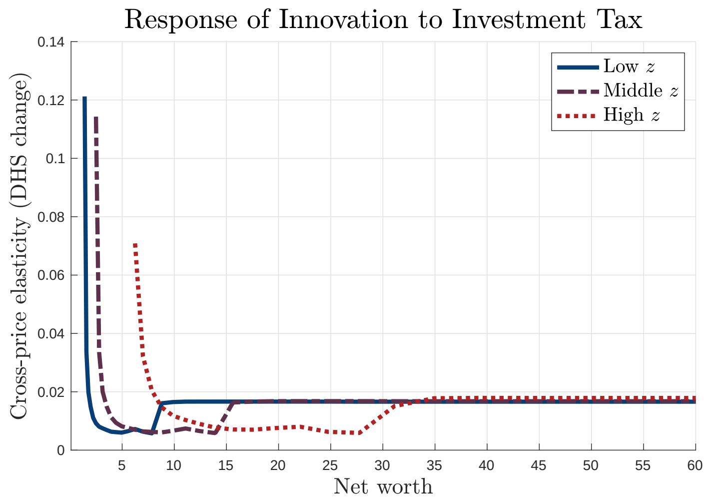
$$\frac{x_{jt}}{k_{jt}} = \alpha_i + \alpha_t + \gamma\frac{1-\tau\zeta_{st}}{1-\tau} + \Gamma' X_{jt} + \epsilon_{jt}, \tag{43}$$

where τ is the corporate tax rate, α_i is a firm fixed effect, α_t is a time fixed effect, X_{jt} controls for cash flows to lagged capital, and ϵ_{jt} are residuals.

To replicate this experiment in our model, we feed in an exogenous shock to the relative price of investment. Appendix B 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 according to an AR(1) with an annual autocorrelation coefficient of 0.8, which implies a half-life around two years (broadly in line with the data). We then simulate a panel of firms from our model's Compustat sample and estimate the regression equation (43). In this regression, we assume all firms face the same present value of tax deductions ζ_t, i.e. there is no sectoral heterogeneity. Since the empirical specification (44) includes time fixed effects to absorb general equilibrium effects, we keep the real interest rate fixed at its initial value $r_t = r^*$ for this exercise. We do not include controls X_{jt} that are outside of our model.

As a reality check, the first two columns of Table E.2 show that the model roughly matches the response of investment to the Bonus Depreciation Allowance. Column (1) shows that the empirical estimate of the regression coefficient is $\widehat{\gamma} = -1.37$, which is close to Zwick

FIGURE E.6: Heterogeneous Responses to the Bonus Depreciation Allowance



Notes: cross-price elasticity of innovation expenditures to the relative price of investment, using 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 0.8 (giving a half-life of roughly two years). Elasticities computed as the Davis, Haltiwanger and Schuh (1998) growth rate in the impact period of the shock.

and Mahon (2017)'s estimate of -1.53 using firm-level IRS microdata. A 50% bonus would increase the average value of $\frac{1-\tau\zeta_{st}}{1-\tau}$ by -0.03, implying its direct effect increased the average firm's investment rate by $-0.03 \times -1.37 = 0.04$, compared to its unconditional average of 0.14. The model's implied regression coefficient in Column (2) is $\widehat{\gamma} = -1.73$, around two standard errors of the empirical estimate.

Column (3) in Table E.2 documents a new empirical finding: the Bonus also substantially raises innovation expenditures. We estimate the regression

$$\frac{\mathrm{RD}_{jt}}{\widetilde{y}_{jt}} = \alpha_i + \alpha_t + \gamma \frac{1-\tau\widehat{\zeta}_{st}}{1-\tau} + \Gamma' X_{jt} + \epsilon_{jt}, \tag{44}$$

which replaces the investment rate on the LHS of (43) with the RD-to-sales ratio $\mathrm{RD}_{jt}/\widetilde{y}_{jt}$. 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) in Table E.2 shows that the model matches the empirical response of inno-

Table E.3

Bonus Depreciation Allowance by Size

	Small firms		Large firms	
	$\frac{\mathrm{RD}_{jt}}{y_{jt}}$, data	(2) $\frac{\mathrm{RD}_{jt}}{y_{jt}}$, model	(3) $\frac{\mathrm{RD}_{jt}}{y_{jt}}$, data	(4) $\frac{\mathrm{RD}_{jt}}{y_{jt}}$, model
$\frac{1-\tau\zeta_{st}}{1-\tau}$	-0.27 (0.10)	-0.27	-0.10 (0.04)	-0.07
R^2	0.83		0.90	

Notes: estimates of $\widehat{\gamma}$ from the regression (43) in columns (1) and (2) or from the regression (44) in columns (3) - (6). Standard errors, reported in parentheses, are clustered by firms. "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 "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 0.8 (giving a half-life of roughly two years).

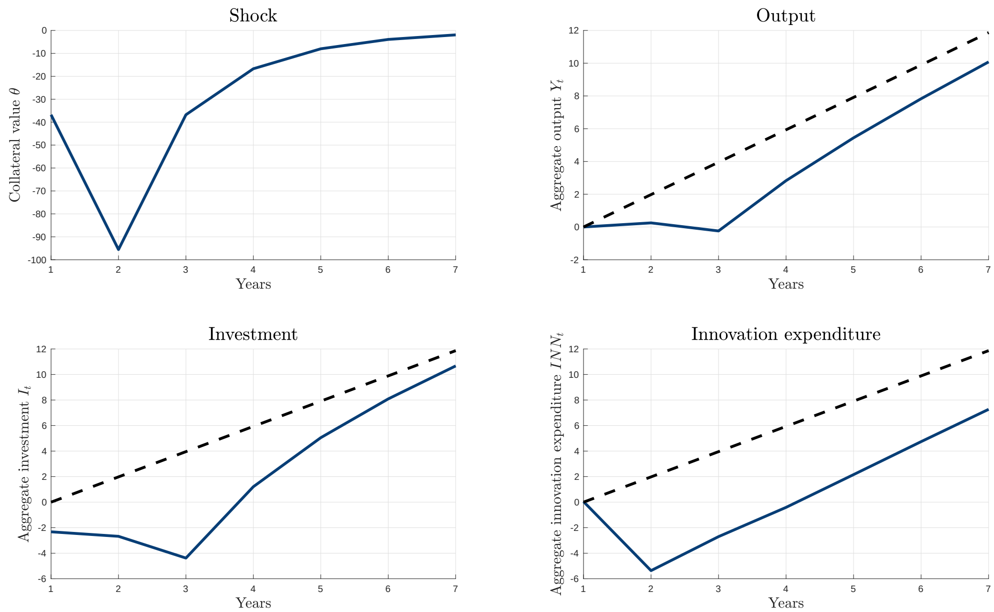
vation to the Bonus within one standard error. In order to understand the role of financial frictions in driving the model's success, Figure E.6 plots the model's cross-price elasticity of innovation with respect to investment. Unconstrained firms have a positive elasticity because higher investment 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 for most constrained firms.

Table E.2 confirms that these size-dependent responses are consistent with 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 four times as responsive to the bonus as are large firms, consistent with our model.

D.4 Transitory Growth Effects of Financial Shocks

We model a financial shock as a transitory decline in the collateral constraint θ_t plotted in the top left panel of Figure E.7. We compute the effects of this shock assuming that the shock is completely unexpected at time $t = 0$ but then agents have perfect foresight as the

Figure E.7: Transition Paths Following Financial Shock θ_t



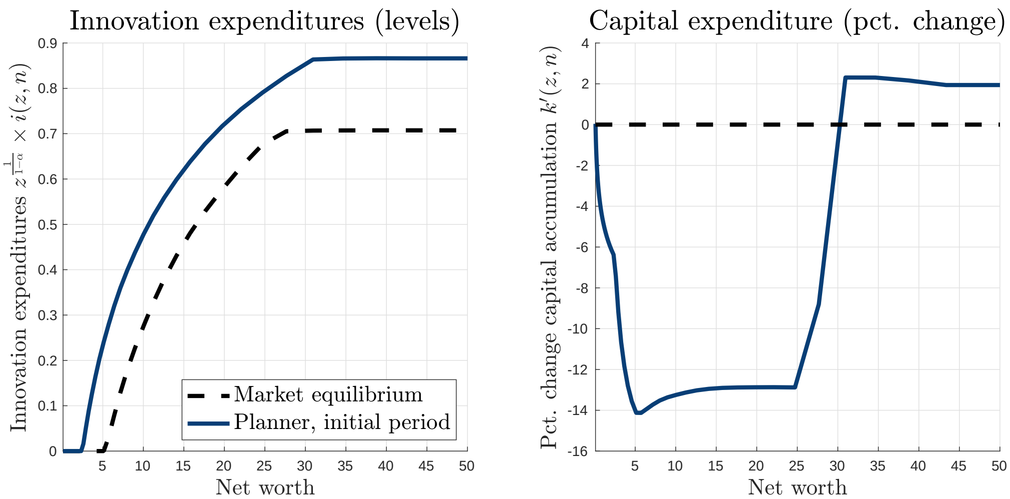
Notes: aggregate transition paths following an unexpected tightening of the collateral constraint θ_t. Top left panel plots the path of θ_t. Remaining panels plot aggregate output, investment, and innovation expenditures in log-deviations from initial period. Dashed black lines are the growth trajectory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

economy transitions back to a BGP. The bottom panels show that the shock reduces both investment and innovation expenditures. However, once the shock has dissipated, the growth rate of investment, innovation, and output return to their original levels. Hence, our model does not generate much internal propagation of financial shocks on aggregate growth rates (although the levels never return to the original trend).

D.5 Details on Planner's Allocation

Figure E.8 compares the equilibrium allocation to the planner's allocation in the first period of the transition path. The left panel shows that the planner increases innovation expenditures for most firms because the shadow of the innovation externality is positive $\Lambda_1 > 0$. The right panel shows that, given this desired increase in innovation, the planner must reduce capital accumulation $k'(z, n)$ by nearly 14% for constrained firms but can increase capital

Figure E.8: Firm-Level Allocations Chosen by Planner



Notes: decision rules in market equilibrium vs. constrained-efficient allocation in initial period of the transition, for a given level of productivity z. Left panel plots innovation expenditures $(A_t z)^{\frac{1}{1-\alpha}} i(z,n)$ for a firm with average productivity z for a given level of net worth. Dashed black line is the private policy rule in the market equilibrium and solid blue line is the planner's policy rule. Right panel plots the percentage difference between the planner's capital accumulation policy relative to the private policy in the market equilibrium.

accumulation around 2% for unconstrained firms.[47]

Hence, as described in the main text, investment and innovation are substitutes for constrained firms but complements for unconstrained firms. In the early stages of the transition path plotted in Figure 8, the substitutability dominates in the aggregate 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 innovation externality Λ_t as described above. Over time, 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.

[47]Note that these effects on flow investment would be approximately $\frac{1}{\delta+g} \approx 10$ times larger than the reduction in capital accumulation plotted in the figure.