

# Knowledge constraints and firm growth

Robert Parham\*

January 3, 2017

## Abstract

Firms are constrained in growing knowledge capital due to frictions on the transfer of firm-specific knowledge (such as the process of recruiting and training skilled labor). Constrained firms must therefore forgo some positive-NPV projects during their growth. I embed micro-founded knowledge constraints in a q-theory model of a firm with two capital goods, structurally estimate the model, and compare the impact of knowledge constraints with that of financial frictions on firm investment and growth dynamics. I find that knowledge constraints subsume the role of financing frictions in explaining the dynamics of R&D-performing firms. These constraints endogenously yield that some firms have higher returns to R&D investment than to physical investment, with estimated differences within the range found in previous empirical work. Small, growing firms are especially affected by the constraint and consequently are more R&D intensive and have higher returns to R&D than larger firms.

*JEL classifications:* D2, G31, G32, J24, O32

---

<sup>0</sup>I thank Ron Kaniel (my adviser), Alex Edmans, Olga Itenberg, John Long, Bill Schwert, Erin Smith, Jerry Warner, Toni Whited, and seminar participants at the University of Rochester for helpful ideas and comments.

\*University of Rochester (robert.parham@simon.rochester.edu, <http://kn.owled.ge>)

# 1. Introduction

The relation between firm-specific knowledge, human capital, and firm growth has been the subject of much theoretical and empirical research. In their discussion of organization capital, Prescott and Visscher (1980) highlight the role of firm-specific human capital, and of the costs of transferring such capital between individuals and firms, in determining firm growth rates. Similarly, Jensen and Meckling (1995) discuss the impact of the cost of knowledge transfer on the spectrum between specific and general knowledge, making the point that “The more costly knowledge is to transfer, the more specific it is.”

This work begins with the assumption that firm-specific knowledge is embodied in the firm’s skilled-labor force. The process of growing the skilled-labor force requires that existing skilled employees transfer their firm-specific knowledge to new-hires in the costly process of recruiting and training. This process makes the current stock of skilled-labor a scarce resource in the “production” of new firm-trained skilled-labor. Firms that find it profitable to grow quickly are hence dynamically constrained and must forgo some positive-NPV investments (i.e., projects become mutually exclusive due to the *internal* scarcity of skilled labor).

Previous research pointed out that firms may be dynamically constrained as a result of human capital frictions. In her seminal work on firm growth, Penrose (1959) writes: “*There are important administrative restraints on the speed of the firm’s growth. Human resources required for the management of change are tied to the individual firm and so are internally scarce. Expansion requires the recruitment of more such resources. New recruits cannot become fully productive overnight. The growth process is, therefore, dynamically constrained.*” While knowledge constraints may somewhat resemble the “time to build” argument for physical capital proposed by Kydland and Prescott (1982), the two are distinct. “Time to build” constraints mean the firm can invest as much as it finds profitable this period (i.e. it does not have to forgo any projects), but the fruits of the investment will be realized with a lag. Knowledge constraints imply the firm must sometimes forgo profitable investment

opportunities due to internal capacity constraints.

My contribution is to formalize this intuition and include it in a standard q-theory model of a firm with two factors of production, physical and knowledge capital. The model also includes convex capital adjustment costs for both capital types, and financial frictions on external capital. I estimate the model using data on skilled-labor-intensive firms by concentrating on R&D-performing firms<sup>1</sup>, and compare the relative importance of knowledge constraints vs. financial frictions in explaining the dynamics of firm investment and growth.

While there is mounting evidence that financial frictions are important in explaining physical investment decisions<sup>2</sup>, the evidence on the impact of financial frictions on R&D investment is mixed, especially for publicly traded firms (Hall and Lerner, 2010). My main finding is that financial frictions are important in explaining firm investment and growth dynamics for my sample of public R&D-performing US firms when knowledge constraints are absent, but lose significance when the model accounts for knowledge constraints. This result corresponds with findings of a recent CFO survey by Jagannathan, Matsa, Meier, and Tarhan (2016) that firms choose to forgo profitable projects because of organizational constraints, and more specifically limited qualified manpower, more than because of financial constraints.

The introduction of knowledge constraints bears on several other salient differences between physical and R&D investment. First, a large body of literature discusses the high returns to R&D. The survey of this literature by Hall, Mairesse, and Mohnen (2010) concludes that on average, the marginal return to the firm on R&D investment is between 2.5 – 4 times higher than on physical investment. Second, as Akcigit (2010) and Akcigit and Kerr (2010) document, there is a negative relation between firm size and innovation intensity—small firms invest more in R&D than do large firms, relative to their respective size. Third, R&D investment is considerably more persistent than physical investment. The

---

<sup>1</sup>The lion's share of R&D investment is salary payments to skilled labor (see OECD, 2012).

<sup>2</sup>E.g., Whited (1992), Cooley and Quadrini (2001), Gomes (2001), Rauh (2006), Chava and Roberts (2008).

$AR(1)$  coefficient for R&D intensity is about 50% higher than the  $AR(1)$  coefficient for physical investment intensity.

Knowledge constraints explain the high returns to R&D by imposing a capacity constraint on investment in R&D because of limits to the growth rate of skilled labor (here, R&D researchers). It is an established result in capital budgeting that in the presence of a capacity constraint, the firm will choose the projects that maximize its NPV and forgo all other projects. The marginal return to knowledge investment in the estimated model is 0.217, more than double the return to physical capital, and within the estimates of returns to R&D described in the survey by Hall et al. (2010).

The negative relation between firm size and innovation intensity is also captured by the model with knowledge constraints. Small firms are more likely than large firms to become constrained in the future following a positive productivity shock, and therefore have a stronger incentive to invest in R&D today. I estimate the coefficient measuring the negative relation between firm size and R&D intensity, and find that it goes from  $-0.025$  in a model without knowledge constraints to  $-0.170$  for a model with knowledge constraints (compared with  $-0.167$  in the data). The persistence of R&D investment is explained by firms predicting future states of the world in which they may be constrained, and smoothing R&D investment to minimize the loss of future positive-NPV projects. The estimated persistence of R&D intensity rises from 0.841 in a model without knowledge constraints to 0.937 in a model with knowledge constraints (compared with 0.931 in the data).

An implication of the model, supported by the data, is that small firms have on average higher marginal returns to R&D than large firms, as they are more likely to be constrained. The model endogenously generates two types of firms—constrained and unconstrained—with constrained firms having higher returns to R&D than unconstrained firms. Instead of exogenously assuming firms of different types based on size, with different returns to R&D, as in Acemoglu, Akcigit, Bloom, and Kerr (2016), the knowledge constraint generates these different type firms endogenously. Different policy recommendations arise from the two models,

as in the Acemoglu et al. (2016) model, high-type firms will benefit from further capital infusion (such as tax credits on R&D investment), whereas my model implies that capital infusion “earmarked” for R&D will not help alleviate the binding knowledge constraints. Rather, firms can alleviate the knowledge constraints by buying entire R&D teams via the “acqui-hire”<sup>3</sup> channel (e.g., Phillips and Zhdanov, 2011; Sevilir and Tian, 2011).

A second implication of the model is that decreasing the friction associated with recruiting and training will increase several observable firm ratios, such as the R&D-to-capital ratio. One way of decreasing the friction is by increasing the training level of the labor pool and thus lowering the training required upon entering the firm. To test this prediction, I construct a proxy of state-level labor-force training that is based on the number of graduating STEM PhDs relative to population. I find that observed variations in firm ratios with respect to the proxy correspond well to implied elasticities derived from the estimated model.

The paper proceeds as follows: the next section presents the model; section 3 discusses the data used; section 4 presents details of the estimation method, moments used, and identification; section 5 discusses the results; section 6 concludes.

## 2. The model

The model presented here is a discrete-time, infinite-horizon model of a firm that produces output using two capital goods: physical and knowledge capital. It follows the knowledge-in-production models originating with Griliches (1979) and the multiple-capital-good q-theory literature (e.g., Hall and Hayashi, 1989; Hayashi and Inoue, 1991; Wildasin, 1984). Hall and Hayashi (1989) provide the first unified treatment of physical and knowledge investment within the same (q-theoretic) framework using knowledge in production. They specify a model similar to the one presented here, but with no adjustment costs, financial frictions, or

---

<sup>3</sup>The process of acquiring a company to recruit its employees, without necessarily showing an interest in its products and services, or their continued operation. Microsoft, for example, made it a business practice throughout the 90s to buy small R&D companies, abolish the product line and reassign the development team to pending projects.

training constraints. Furthermore, as they lack a method to estimate models with no closed form first order conditions, and due to the unobservable nature of the knowledge stock, they are forced to make several assumptions they call “heroic” to reach estimable equations.

Following the popularization of structural estimation and indirect inference<sup>4</sup> methods in the financial literature, which allow estimation of models with no closed form first order conditions and with unobservable state variables, several authors have revisited the classical Hall and Hayashi (1989) model (e.g., Li and Liu, 2012; Warusawitharana, 2014). These authors do not discuss the human-capital nature of R&D investment or consider a knowledge constraint mechanism similar to the one presented here. My principal deviation from existing R&D models is in the way I model the law of motion for knowledge capital—capturing the training and recruiting frictions and the knowledge constraints. In this, I heed the call of Hall et al. (2010), who conclude that “because the additive model is not really a very good description of knowledge production, further work on the best way to model the R&D input would be extremely desirable.”

### 2.1. *The firm*

At the beginning of each period, the firm observes the levels of physical capital  $K_t$ , knowledge capital  $N_t$ , and the realization of its persistent productivity  $Z_t$ . The firm combines the two types of capital into a composite capital with elasticity of physical capital  $0 < \omega < 1$ , and then uses the composite capital in producing cashflow (net of the cost of goods sold and payments to unskilled labor) with elasticity  $0 < \theta < 1$ , such that:

$$CF_t = e^{Z_t} (K_t^\omega N_t^{1-\omega})^\theta \tag{1}$$

The firm’s productivity,  $Z_t$ , follows a standard AR(1) process with persistence  $\rho_Z$  and

---

<sup>4</sup>See Gourieroux, Monfort, and Renault (1993) and Gourieroux and Monfort (1996).

mean  $\mu_Z$ , with  $\mu_Z$  capturing the average total factor productivity (TFP) across firms:

$$Z_{t+1} = (1 - \rho_Z)\mu_Z + \rho_Z Z_t + \epsilon_{Z,t}, \quad (2)$$

and with the innovation term,  $\epsilon_{Z,t}$ , normal i.i.d. with mean zero and standard deviation  $\sigma_Z$ :

$$\epsilon_{Z,t} \sim \mathbb{N}[0, \sigma_Z]$$

After the cashflow is realized, the firm chooses its level of investment in physical capital,  $IK_t \geq 0$ , and investment in knowledge capital,  $IN_t \geq 0$ . Once these choices are made, the firm's earnings are realized and given by

$$E_t = CF_t - IK_t - IN_t \quad (3)$$

## 2.2. Financial frictions

If the firm chooses to invest more than it has generated in cashflow (i.e., if  $E_t < 0$ ), it must raise costly external capital to cover this financing gap. The model abstracts from capital structure, so external capital is an unspecified combination of debt and equity. As in previous models of firm dynamics, such as Hennessy and Whited (2005) and DeAngelo, DeAngelo, and Whited (2011), I assume the cost of raising external capital is convex. There is strong empirical support for this assumption for both equity and debt financing.<sup>5</sup> Specifically, I follow Hennessy, Levy, and Whited (2007) in defining the firm's distributions to (or, when negative, capital infusions from) stakeholders, denoted  $D_t$ , as:

$$D_t = E_t - \Phi[D_t, K_t], \quad (4)$$

---

<sup>5</sup>Krasker (1986), Asquith and Mullins (1986), Fazzari, Hubbard, and Petersen (1988), Altinkilic and Hansen (2000), Weiss (1990).

with:

$$\Phi[D_t, K_t] = \mathbf{1}[D_t < 0] \frac{\gamma}{2} \left( \frac{D_t}{K_t} \right)^2 K_t, \quad (5)$$

in which  $\mathbf{1}[D_t < 0]$  is an indicator that takes the value 0 if  $D_t \geq 0$  and 1 if  $D_t < 0$ . The parameter  $\gamma$  captures the magnitude of the external capital costs. Under this formulation, if  $E_t \geq 0$  then there are no external capital costs ( $D_t = E_t$ ). If  $E_t < 0$ , the firm bears the cost of raising external capital on the entire amount it raises, which includes both  $E_t$  and the flotation cost itself.

### 2.3. *Physical investment*

The law of motion for physical capital is:

$$K_{t+1} = (1 - \delta_K)K_t + \Lambda^K \left[ \frac{IK_t}{K_t} \right] K_t \quad (6)$$

in which  $\delta_K$  is the depreciation rate of physical capital. Here, I follow Uzawa (1965) and Hayashi (1982) in formulating the adjustment costs to physical capital as affecting the accumulation equation for physical capital. Specifically, I use the functional form proposed by Jermann (1998) and used in the financial literature by Kung and Schmid (2015):

$$\Lambda^K \left[ \frac{IK_t}{K_t} \right] = \hat{\lambda}_{1K} + \hat{\lambda}_{2K} \left( \frac{IK_t}{K_t} \right)^{\lambda_K} \quad (7)$$

in which  $\lambda_K$  captures the (inverse) elasticity of the investment rate.<sup>6</sup> As in Kung and Schmid (2015), I set  $\hat{\lambda}_{1K}$ ,  $\hat{\lambda}_{2K}$  such that there are no adjustment costs in the deterministic steady state, when firms just replenish depreciated capital,  $IK_t = \delta_K K_t$  (by requiring  $\Lambda^K[\delta_K] = \delta_K$ , and its first derivative,  $\Lambda_1^K[\delta_K] = 1$ ).

---

<sup>6</sup>Such that lower  $\lambda_K$  values imply higher friction.

## 2.4. Knowledge investment and the firm academy

Firms may be categorized into one of three possible states, with respect to knowledge investment:

1. Shrinking or steady ( $N_{t+1} \leq N_t$ )
2. Growing but unconstrained ( $N_t < N_{t+1} < N^{max}$ )
3. Constrained ( $N_{t+1} = N^{max}$ )

Shrinking or steady firms are assumed to be unaffected by the training and recruiting friction<sup>7</sup>, and their law of motion for knowledge capital is similar to the one for physical capital:

$$N_{t+1} = (1 - \delta_N)N_t + \Lambda^N \left[ \frac{IN_t}{N_t} \right] N_t \quad (8)$$

$$\Lambda^N \left[ \frac{IN_t}{N_t} \right] = \hat{\lambda}_{1N} + \hat{\lambda}_{2N} \left( \frac{IN_t}{N_t} \right)^{\lambda_N} \quad (9)$$

in which  $\delta_N$  is the depreciation rate of knowledge capital. An important theoretical point when discussing knowledge depreciation is that being an information good, knowledge does not depreciate in the usual sense. Knowledge depreciation is “not physical forgetting but rather the dissipation of rents as a result of obsolescence” (Griliches (1998), p. 271), following the creative destruction concept of Schumpeter (1942). The (inverse) elasticity of the investment rate is captured by  $\lambda_N$ , and  $\hat{\lambda}_{1N}$ ,  $\hat{\lambda}_{2N}$  are again set such that there are no adjustment costs when the firm maintains its level of knowledge capital, i.e., when  $IN_t = \delta_N N_t$ . The equation holds for shrinking or steady firms, for which  $IN_t \leq \delta_N N_t$ .

For  $IN_t > \delta_N N_t$ , I assume that some of the firm’s current researchers are set aside as instructors in a “firm academy”, hiring and training new personnel. There is no substitutability between instructors and new-hires, and so the production function of the academy is fixed-proportions (Leontief). The firm academy captures existing entities in high growth

---

<sup>7</sup>A full labor model will take account of exogenous separation of part of the skilled labor force every period and the need to train replacements, but this consideration is likely not a first-order issue for the purpose of the current discussion.

firms, such as Facebook’s “Bootcamp” and Twitter’s “University”, which are aimed at reducing the friction associated with hiring and training new skilled labor.<sup>8</sup> While many firms may not have a firm academy and rely on the general pool of researchers for interviewing and on-the-job training, a division-of-labor argument implies the firm academy model represents a lower bound to the friction that hiring and training imposes on current R&D researchers.

We can decompose the payment to skilled labor into three parts as follows:

$$IN = \overbrace{(IN - \delta_N N)}^{\text{new recruits}} + \overbrace{\frac{1}{\alpha}(IN - \delta_N N)}^{\text{instructors}} + \overbrace{(\delta_N N - \frac{1}{\alpha}(IN - \delta_N N))}^{\text{researchers}} \quad (10)$$

with the first term the payment to new skilled labor that are recruited this period and entering the academy to be trained, the second term the payment to current skilled labor diverted to be instructors at the academy (with each instructor training  $\alpha$  new recruits), and the remaining trained, skilled labor continuing their work as R&D researchers.

By requiring payment to researchers to be non-negative, I have the academy constraint:

$$\delta_N N_t - \frac{1}{\alpha}(IN_t - \delta_N N_t) \geq 0 \quad (11)$$

or

$$IN_t \leq (1 + \alpha)\delta_N N_t \quad (12)$$

## 2.5. Firm value

Firm value at time  $t$  can now be expressed using the bellman equation:

$$V[K_t, N_t, Z_t] = \max_{IK_t, IN_t} \left\{ D_t + \frac{1}{1+r} \mathbb{E}_{Z_{t+1}/Z_t} V[K_{t+1}, N_{t+1}, Z_{t+1}] \right\} \quad (13)$$

---

<sup>8</sup>Another example, which was the original inspiration for this idea, is the Israeli Defense Force’s cyber-security academy, ARAM.

subject to the laws of motion of for  $K_t$  and  $N_t$  which were specified above, and with  $r$  the expected rate of return (weighted average cost of capital).

### 3. Data and stylized facts

I estimate the model using data obtained from the Compustat-CRSP Combined dataset. For each firm-year observation, the data include all accounting items from the firms' end of year balance sheets and cashflow reports, as well as the market value of the firms' equity at the end of the year. Firm-year observations omitted are those: with missing values for important accounting items; missing two lags; which have less than \$1M in physical capital or market value; or are in the financial or utilities industries.<sup>9</sup> As I concentrate on R&D-performing firms, I drop all firm-years in which zero or missing R&D is reported. The final data are an unbalanced panel of 35,589 firm-year observations on a set of 4,504 distinct publicly traded R&D-performing firms between the years 1983 and 2014. The cross-section varies in size between 975 firms in 1983 and 1,337 firms in 1997.

Table 1 defines the main variables used in the analysis in terms of Compustat variables. Firm value ( $V$ ) is defined as the market value of equity plus the book value of debt. Accordingly, distributions ( $D$ ) are defined as the distributions to equity holders (dividends and stock repurchases minus new stock issuance) plus distributions to bond holders (interest minus increase in book value of debt), following Bradshaw, Richardson, and Sloan (2006). These definitions are made so I can abstract from the firm's capital structure and dividend choices. It is worth noting that as accounting rules dictate that R&D be expensed rather than capitalized, it is a component of Selling, General and Administrative expenses (SG&A). The definition of cashflow from operations ( $CF$ ) is therefore adjusted accordingly. Also note that while physical capital ( $K$ ) is observable in the data, knowledge capital ( $N$ ) and productivity ( $Z$ ) are used as state variables in the model but have no observable counterpart in the

---

<sup>9</sup>These industries were dropped as the interpretation of their balance sheet items is significantly different from that of the firms that remain in the sample.

data.

One challenge when fitting the model to data is the treatment of firm acquisitions. The model does not include a concept of acquisitions, but they are a way the firm can increase its labor force, knowledge capital, and physical capital. Three possible approaches to the treatment of acquisitions are to: drop firm-years with significant acquisition activity, or impute acquisitions into firm investments in physical and knowledge capital, or ignore them altogether. In my baseline estimation, I opt for the second approach—imputing acquisitions into investment—though I verify my conclusions are not affected by using the two other approaches. Acquisitions are split between the two investment streams by the ratio of investments in both capital types prior to the acquisition. I also verify that using an index constructed from observed increases in physical property and goodwill at the acquiring firm, post-acquisition, does not change my conclusions.

The first notable (and well known) stylized fact about firm variables is that they are exponentially distributed. It is therefore common to consider them in terms of intensity relative to firm size. I use firm physical capital ( $K$ ) as the firm size deflator, as it is observable and directly maps into one of the state variables of the model. I also verify that using firm value ( $V$ ) as the deflator leads to similar conclusions. A second notable fact is that even after deflating by firm size, the intensities of positive values still exhibit an exponential distribution, as is evident in Appendix Fig. A.1, which presents the distributions of the *logs* of physical and knowledge investment intensities in the data. The logs of the intensities are approximately normally distributed. A similar pattern holds for the other intensity variables. I therefore apply a log-transform to all the non-negative intensities, as the moments of exponential distributions are strongly affected by tail values and are poorly behaved. Finally, I reduce the heterogeneity of the data by taking industry and year fixed effects before calculating the reported moments. Industries are based on the Fama and French (1997) categorization.

The average firm in the sample of R&D-performing firms has \$167M in physical capital,

and has a physical investment intensity of 11.8% of physical capital, compared with 14.4% of physical capital invested in R&D. Firms in the sample are more R&D-intensive than physical-capital-intensive, which is driven by the fact all firms that do not perform R&D were dropped from the sample. This is a strong indication of the dependence of these firms on skilled-labor in their production and growth.

The dynamic behavior of knowledge and physical investment is markedly different. While cashflow intensity has a persistence of 0.876, R&D intensity is more persistent at 0.931, and physical intensity less persistent at 0.583. As documented by Akcigit (2010) and Akcigit and Kerr (2010), R&D and physical investment intensities also differ in the way they scale with firm size. While both intensities decrease with firm size, the decrease in knowledge investment intensity is more than five times higher than that of physical investment intensity - larger firms are much less likely to invest in R&D than they are to invest in physical assets. Importantly, this result holds when using firm value rather than firm physical capital as the measure of firm size.

In summary, R&D-performing firms invest more in knowledge than physical capital, and change knowledge investment intensity slower over time. Furthermore, small firms invest more in both knowledge and physical capital relative to their size, but the decrease in R&D investment as they grow is faster than the decrease in physical investment. Finally, as documented by previous research, the return to investment in knowledge capital is considerably higher than the return to investment in physical capital, and this difference is higher for smaller firms (Acemoglu et al., 2016). The next section shows that the standard q-theory model with investment frictions and convex adjustment costs is unable to explain these stylized facts, and shows that accounting for recruiting and training frictions allows the model to explain these facts.

## 4. Estimation

This section describes the process of taking the model to data. As the model lacks closed form first order conditions that can be directly estimated, and includes unobservable state variables such as the stock of knowledge capital, I estimate it using the simulated method of moments (SMM), an indirect inference method (Gourieroux et al., 1993; Gourieroux and Monfort, 1996). This section also describes the moments targeted in the estimation and provides economic intuition regarding the identification of the model parameters by these moments. It concludes with several technical details regarding the estimation, such as the imposition of a tax regime on the model, and the addition of exit and entry dynamics. The Appendix presents identification tests, sensitivity analysis, and the response of model moments to the parameters.

### 4.1. Estimation method

To solve the model, I numerically iterate on the model’s Bellman equation (Eq. 13).<sup>10</sup> For a given parameter vector  $b$

$$b = \{\theta, \omega, \rho_Z, \mu_Z, \sigma_Z, \delta_K, \lambda_K, \delta_N, \lambda_N, \alpha, \gamma, r\}, \quad (14)$$

this process yields a value function  $V[K_t, N_t, Z_t]$  and a policy function  $\{K_{t+1}, N_{t+1}\} = H[K_t, N_t, Z_t]$ .

With the value and policy functions for a given parameter vector  $b$  in hand, I generate  $S = 10$  simulated panels, each consisting of 5,000 firms, over 130 periods. I drop the first 100 periods of each simulation, to allow the distribution of firm productivity,  $Z$ , to stabilize, and keep the last 30 periods. For each of the 10 simulated panels, I calculate a vector of moments of the simulated data, such as mean, persistence and standard deviation of various firm ratios and growth rates. The choice of moments is discussed in the next section. Let

---

<sup>10</sup>The exposition in this section follows the discussion in DeAngelo et al. (2011).

$\mathbb{M}[\hat{Y}^s(b)]$  denote the vector of moments generated from the simulated data  $\hat{Y}^s$ , for a given parameter  $b$ , and let  $Y$  denote the Compustat-CRSP data described in the previous section. Next, define

$$G[b] = \left\{ \mathbb{M}[Y] - \frac{1}{S} \sum_{s=1}^S \mathbb{M}[\hat{Y}^s(b)] \right\}, \quad (15)$$

to be the vector of differences between the simulated and data moments.

The SMM estimator of the parameter vector  $b$ , denoted  $\hat{b}$ , can now be defined as

$$\hat{b} = \arg \min_b G[b]'WG[b], \quad (16)$$

in which  $W$  is a positive-definite weighting matrix. I use the inverse of the sample variance-covariance matrix of the moments used, which I calculate using the influence-function approach described by Erickson and Whited (2000).

The asymptotic variance-covariance of the SMM estimator  $\hat{b}$  is given by

$$avar(\hat{b}) = \left(1 + \frac{1}{S}\right) \left[ \frac{\partial G(b)}{\partial b} W \frac{\partial G(b)}{\partial b'} \right]^{-1} \left[ \frac{\partial G(b)}{\partial b} W \Omega W \frac{\partial G(b)}{\partial b'} \right] \left[ \frac{\partial G(b)}{\partial b} W \frac{\partial G(b)}{\partial b'} \right], \quad (17)$$

in which  $\Omega$  is the sample variance-covariance matrix of the moments used (i.e.,  $\Omega = W^{-1}$ ).

Finally, the overidentifying restrictions of the model can be tested using the  $J$ -statistic

$$J = \frac{S}{S+1} G[\hat{b}]'WG[\hat{b}] \quad (18)$$

that converges in distribution to a  $\chi^2$  with degrees of freedom equal to the number of moments minus the number of parameters.

I conduct a search on the parameter space to find the  $\hat{b}$  that minimizes the  $J$ -statistic by using a multi-start algorithm to search for starting points, and a non-derivative-based pattern-search algorithm to optimize from the chosen starting points. The use of non-derivative optimization is important because the knowledge constraint introduces kinks into the policy function and so the search space may be non-differentiable.

## 4.2. Moments and identification

For the model to be identified, the moments used in the estimation need to be responsive to the parameters used. While SMM identifies all parameters using all moments, some moments are nevertheless more informative regarding some parameters. This subsection describes the moments being used, and the economic intuition behind the parameter identification. Note that a crucial difficulty with identification is the unobservable nature of the stock of knowledge capital,  $N$ .

The first column of Table 2 describes the moments used in the estimation. For convenience of discussion, they are split into four groups. The first group is moments that pertain to firm size (defined as the log of physical capital,  $K$ ). I target the mean and standard deviation of firm size in the data, as well as the mean and standard deviation of the growth rate of firm size. The next group is moments that pertain to the activity of the firm, such as cashflow, distributions and firm value. Here, I target the mean log value-to-physical-capital ratio, as well as the standard deviation of its growth rate; the mean cashflow-to-capital ratio and its persistence; and the mean and standard deviation of the ratio of distributions-to-value. The third group contains moments related to firm R&D investment. I target the mean log R&D-to-capital ratio along with its persistence and the standard deviation of its growth rate; the standard deviation of log R&D investment, and the slope coefficient from regressing log R&D-to-capital ratio on firm size. The last group is of moments related to firm physical investment. Here, I target the mean log investment-to-capital ratio along with its persistence and the standard deviation of its growth rate; the standard deviation of log physical investment, and the slope coefficient from regressing log investment-to-capital ratio on firm size.

The parameters affecting firm size and cashflow,  $\theta, \rho_Z, \mu_Z, \sigma_Z$  are fairly readily identified. The return to scale parameter  $\theta$  determines the optimal scale of operation. The mean firm physical size ( $\log K$ ) is very sensitive to this parameter and pins it down well. The mean of the productivity process  $\mu_z$ , which captures total factor productivity, controls the rate by

which composite capital is transformed into gross profits, and so is pinned down by the mean cashflow-to-capital ratio,  $(CF/K)$ . The persistence of the productivity process  $\rho_Z$  directly affects the AR(1) coefficient of the cashflow-to-capital ratio  $(CF/K)$ . The parameter governing the magnitude of innovation to the shock process,  $\sigma_Z$ , determines the standard deviation of the growth rate of cashflow-to-capital  $(CF/K)$ . Finally, the cross-sectional distribution of firm sizes is affected by all four parameters, and so the standard deviation of firm size ( $\log K$ ) and the standard deviation of its growth rate are added as control moments.

Given a pin-down of the productivity process, which controls the “growth options” available to firms, the elasticity of physical capital in composite capital production,  $\omega$ , can now be pinned down. Firm value is composed of the value of physical capital, knowledge capital, and growth options. The optimal ratio of physical capital to knowledge capital, controlled by  $\omega$ , directly affects the mean of log firm value to physical capital  $(V/K)$ . The expected rate of return,  $r$ , is pinned down well by the mean of distributions to value  $(DV/V)$ , along with the mean growth rate of firm size ( $\log K$ ), as the two jointly determine the ex-post rate of return to stakeholders.

The parameters governing firm physical investment behavior are well-identified as well, due to the observability of the stock of physical capital,  $K$ . The rate of physical capital depreciation  $\delta_K$  determines the average intensity of investment in physical capital relative to the stock of physical capital, and so the mean of log investment intensity  $(IK/K)$  captures it well, as in other structural investment models. Adjustment costs on physical capital, governed by  $\lambda_K$ , control how costly it is for firms to deviate from investing just the amount required to cover depreciation. As such, the dispersion in the growth rate of investment intensity  $(IK/K)$  responds to this parameters and identifies it.

Due to the un-observability of the knowledge capital stock, identification of the parameters governing firm knowledge investment behavior is somewhat more tenuous. Via a similar argument to physical capital, the rate of knowledge capital depreciation  $\delta_N$  would optimally be identified by the mean of log R&D-to-knowledge-capital intensity,  $(IN/N)$ . Unfortunately,

$N$  is not available in the data. As the firm aims to maintain a constant optimal ratio between its stock of knowledge capital,  $N$ , and its stock of physical capital,  $K$ , governed by  $\omega$ , I use the log of R&D-to-physical-capital ratio ( $IN/K$ ) to identify  $\delta_N$ . Similarly, I use the standard deviation of the growth rate of R&D intensity with respect to physical capital ( $IN/K$ ) to identify the adjustment costs on knowledge capital, governed by  $\lambda_N$ . I also target the standard deviation of the growth rate of raw R&D, without dividing by  $K$ , to provide additional control to these parameters, as it is affected by both of them but not directly by  $K$ .

An increase in the financial friction parameter,  $\gamma$ , makes all external capital (negative distributions) more expensive, leading the firm to use less of it, and decreasing the dispersion of the ratio of distributions-to-value ( $DV/V$ ). This moment hence identifies the financial friction parameter. A second moment useful in the identification of the financial friction is the dispersion of the growth rate of log value-to-physical-capital, as a high friction means slower response of firm value to change in productivity, and hence slower growth of this ratio.

Finally, we are left with the firm academy parameter  $\alpha$ . The firm academy constrains the ability of the firm to invest in R&D, and this constraint depends on the current stock of R&D researchers. Hence, each existing researcher gains more option value in case of future need for growth. This induces the firm to smooth the investment in R&D, and so affects the persistence of log R&D-to-physical-capital ratio ( $IN/K$ ). Furthermore, as smaller firms are more susceptible to the constraint, their investment intensity increases relative to larger firms. The slope coefficient from a regression of log R&D-to-physical-capital intensity ( $IN/K$ ) on firm size ( $\log K$ ) captures this dependence of intensity on size and helps pin down  $\alpha$  as well.

### 4.3. Technical notes

Prior to estimation, I impose a tax regime on the model, to allow it to capture the observable quantities in the data. Eq. 3 is hence modified to be

$$E_t = (1 - \tau_C)(CF_t - IN_t) - IK_t + \tau_C \delta_K K_t + \tau_N IN_t, \quad (19)$$

with  $\tau_C$  the corporate tax rate, which I set at the statutory rate of 35%, and with  $\tau_N$  the Research & Experimentation Tax Credit established by the Economic Recovery Tax Act of 1981. Internal Revenue Code 41, which codifies the R&D tax credit, defines several methods for calculating the amount of R&D spending eligible for tax credit (e.g., spending above a certain threshold defined by the average of firm R&D spending in the past 4 years). As such, the mechanism could not be easily introduced into the model. I set this rate at 4% following Moris (2005) and Hall, Jaffe, and Trajtenberg (2001).

As the dynamics described by the model are especially relevant to the case of young, small firms with high growth potential, they are closely related to the dynamics of entry and exit, or the *dynamism* of the economy. To achieve such dynamism, I include entry and exit when estimating the model. Firms exit the simulated panel either when their value becomes negative because of bad shocks to their productivity (limited liability), or they become too small (de-listing) or at an exogenous rate  $\pi$ . To maintain a constant measure of firms in the economy, an exiting firm is immediately replaced by an entering firm. Entering firms draw a productivity level from the current productivity distribution of the simulated panel, but begin with minimal capital levels, as in Itenberg (2014).

The rate of firm exit is pinned down by the exponential decay pattern of firm age in the data, presented in Appendix Fig. A.2. Every year, approximately 6.4% of firms exit, and this rate is roughly independent of firm age, as is evident by the linear relation in Fig. A.2. I therefore set  $\pi$  such that total exit in the model (due to low value, small size, and exogenous exit) is 6.4% per period.

## 5. Results

This section presents the main estimation results, comparing the two nested models with and without knowledge constraints. The model lacking knowledge constraints fits the data fairly well, but is unable to match the persistence and the decrease with size of R&D intensity. It is also unable to generate high returns to R&D, or differential returns to R&D with firm size. Adding knowledge constraints allows the model to capture these features of the data. I then present the results of estimating the model separately within each industry, and compare the results with measures of industry training needs and financial frictions. I also present a validation test of the model-implied elasticity of several firm ratios with respect to the strength of the knowledge constraint. Finally, I discuss a prediction of the model regarding firm growth paths and Gibrat's law.

### 5.1. *Main results*

Table 2 presents the data moments and the simulated moments from an estimated model that does not include a knowledge constraint (No Academy:  $\alpha \rightarrow \infty$ ) and an estimated model that includes a knowledge constraint (Academy:  $\alpha$  is estimated). The model without knowledge constraints establishes a baseline and measures the extent to which a standard two-capital q-theory model with financial and capital accumulation frictions is able to capture the dynamics of R&D performing firms. The model with knowledge constraints estimates the added benefit of the knowledge constraint in explaining firm dynamics.

The model without knowledge constraints is able to fit most data moments fairly well. The model overshoots somewhat for average firm size, with larger firms in the simulated model, but is able to capture the other aspects of firm size, like the dispersion of firm size, and the mean and dispersion of firm growth rates. It significantly undershoots the standard deviation of the ratio of distributions to value, with estimated value in the model almost half that of the data. This undershooting is to be expected, as the simulated firms lack a CFO

and the decisions on firm distribution vs. retention of earnings that a CFO would make. Notably, the model is unable to fit either the persistence of R&D intensity, or its decrease with firm size. Finally, the persistence of physical investment intensity in the simulated model is significantly higher than in the data. This is caused by the stark difference between the persistence of the cashflow process and that of physical investment. The overshoot can be remedied by adding a fixed cost of physical capital adjustment to the model, as in DeAngelo et al. (2011), following Cooper and Haltiwanger (2006).<sup>11</sup> The cost of doing so is the addition of another parameter, which requires estimation, to the model.

The column of Table 2 marked “Acad” reports the moments for simulated data drawn from the model *with* an academy. The persistence of R&D intensity increases from 0.841 in the model without an academy to 0.937 in the model with an academy, compared with 0.931 in the data. The scaling coefficient measuring the decrease in R&D intensity with size is also captured by the model, increasing in magnitude by a factor of 7 from  $-0.025$  without knowledge constraints to  $-0.170$  with knowledge constraints, almost precisely the value in the data. The persistence of physical intensity decreases, but it is still higher than the persistence in the data.

The estimated parameter values for both models are reported in Table 3. Note that the firm academy parameter  $\alpha$  is missing in the first model by design, as the firm academy is shut down in the first model. Most parameter estimates are stable between the two models. Four parameters significantly differ between the two:  $\mu_Z$ —capturing average TFP,  $\lambda_N$ —the convex knowledge capital adjustment parameter,  $\alpha$ —the firm academy parameter, and  $\gamma$  - the financial friction parameter.

The fact average TFP is now economically close to zero and statistically insignificant implies the model with an academy allows firm knowledge capital to fully explain the average productivity of firms, which is a pleasing feature of the model. The increase in  $\lambda_N$ , the convex knowledge capital adjustment parameter, means convex costs of adjustment are lower in the

---

<sup>11</sup>TODO - Add fixed adjustment costs in a robustness model.

model with an academy, as there is an inverse relationship between  $\lambda_N$  and the magnitude of the costs of adjustment. It is explained by the addition of a different source of friction—the knowledge constraint. The parameter governing the tightness of the knowledge constraint, *alpha*, is estimated at about 2.5 and is statistically significant. The most intriguing result is that  $\gamma$ , the financial friction parameter, decreases by a factor of 10, and becomes economically and statistically insignificant.

The decrease in the financial friction parameter  $\gamma$  suggests that once accounting for knowledge constraints, financial frictions are not required in order to explain firm dynamics for my sample of public R&D-performing firms. An inability to grow investment quickly is attributed to financial frictions in a model lacking knowledge constraints. This attribution implies a scarcity of capital to be invested in R&D (scarcity of supply of capital). A model including knowledge constraints instead implies a lack of opportunities to invest—not because firms lack profitable opportunities—but rather because they are unable to capture them due to constraints on the growth of skilled labor (scarcity of demand for capital).

This inability to capture profitable opportunities means firms will choose those investment maximizing their NPV (i.e., promising the highest returns), while forgoing all other investments. A simple way to develop intuition regarding this mechanism is to consider a firm with a very high cost of capital. The firm is unconstrained and invests in both physical and knowledge capital up to the point at which the marginal return on investment is equal for both types of investment, and equal to the cost of capital. Reducing the cost of capital now implies the firm has more investment opportunities (at lower returns on investment) and both types of investment should increase. This can continue up to the point at which the firm becomes constrained by knowledge constraints. Further decrease in the cost of capital will not increase knowledge investment, making returns to R&D higher than returns to physical investment, and higher than the cost of capital.

Panels (a) and (b) of Fig. 1 present the model-implied average (across firms of the same size) of the marginal rates of return to investment in knowledge and physical capital, re-

spectively, for the model lacking knowledge constraints. The returns are derived directly from the estimated firm value function, and are graphed with respect to firm size. As the principle of equating marginal cost to marginal benefit dictates, firms invest in both capital types enough such that the net marginal return on investments is equal to the cost of capital ( $r$ , estimated at 9.1%). This is true for firms of all sizes.

Panel (c) and (d) of Fig. 1 repeat this analysis, but for the model with knowledge constraints. The estimated marginal return on physical investment is again equal to the cost of capital for firms of all sizes, as expected in a q-theory model. The pattern of marginal returns on investment in knowledge capital is, however, starkly different. Firms have higher marginal returns to knowledge investment than their cost of capital (or the returns to physical investment), and this is especially accentuated for smaller firms. This finding implies that a further investment in knowledge capital would increase firm value, but firms are unable to take this further investment due to the constraint. As small firms are more susceptible to the constraint, they forgo more projects and have higher marginal returns.

The decreasing intensity of R&D with firm size described earlier interacts with, and is partly explained by, the decreasing marginal return to R&D investment. Smaller firms are more likely to be constrained, which causes their high return to R&D. Relieving the constraint would increase firm value, and requires building a larger knowledge stock. Larger firms are less likely to be constrained, and so have less of a “precautionary saving” motive to investing in R&D. Hence, smaller firms invest more in R&D, relative to size, than larger firms.

## 5.2. *By-industry estimates*

TODO - Add per-industry estimates. Compare financial constraint parameter and knowledge constraint parameter with external measures of financial constraints and firm cost of training, respectively.

### 5.3. Validation

One of the strengths of structural modeling is the ability to conduct counter-factual analysis, by holding all parameters but one constant and testing the effects of a change in this one parameter on observable outcomes of the model. The parameter of interest for such an exercise in the context of knowledge constraints is  $\alpha$ , which measures the strength of the constraint, stemming from the need for within-firm training.

Comparing the results of such a counter-factual analysis with results based on data is challenging, as we require data on variation in within-firm training costs, which is difficult to come by. However, as highlighted by Mincer (1962), “*the same degree of occupational skill can be achieved by shortening formal schooling and lengthening on-the-job training or by the reverse*”. This insight implies that variation in the training level of the labor force available for firms to hire will induce variation in the friction imposed by firm-training.

To that end, I construct a proxy based on the number of STEM<sup>12</sup> PhDs graduating within every state. The proxy is based on data from the Carnegie Classification of Institutions of Higher Education, and roughly follows Kalcheva, McLemore, and Pant (2016). Unlike the proxy by Kalcheva et al. (2016), which counts the number of PhD-granting institutions within each state, I specifically count the number of STEM PhDs, and transform it into a per-capita measure by dividing it by the population of every state. I then use this proxy, along with the reported headquarter state of each firm in my sample, to test for variation in knowledge constraints.

This proxy has several notable shortcomings. First, one of the hallmarks of the US labor market is the high degree of mobility of skilled labor between states. Second, data on graduating STEM PhD is not available as a panel but rather only as a single cross-section in the year 2010. Third, there is no guarantee that the state at which a firm is headquartered is the state at which it conducts most of its R&D. Nevertheless, this proxy provides a first approximation to the degree of training of a state’s labor-force, and to the

---

<sup>12</sup>Science, Technology, Engineering, Math.

knowledge constraints faced by firms.

Table 4 reports the results of this validation analysis. I consider several growth rates and observable firm ratios, and regress each on the labor-skill proxy and a full set of industry and year fixed effects. As firms rarely change the state of their headquarter, I am unable to take firm fixed effects, but standard errors are clustered by firm. For ease of interpretation, the labor-skill proxy is standardized. Many of the firm observables exhibit significant variation with the labor-skill proxy, with the most notable ones the R&D intensity, and the cashflow- and value-to-capital ratios. Firms headquartered in states with better-trained labor force are more R&D intensive, profitable and valuable relative to their counterparts within the same industry located elsewhere.

Model validation stems from comparing these findings to the elasticity of the firm observables to changes in  $\alpha$  implied by the counter-factual analysis within the model. The last column of Table 4 reports the model-implied elasticities with respect to  $\alpha$ , and it is reassuring to see that both signs and magnitudes align fairly well with the results of the data exercise based on the labor-skill proxy.

#### 5.4. *Firm growth paths*

Previous research (e.g., Evans, 1987) has documented the failure of Gibrat’s law, which states that the proportional rate of growth of a firm is independent of its size. On average, a robust feature of the data is that small firms grow faster than large firms, even after controlling for selection bias (attrition). Another way of stating this fact is that firms’ growth paths are on average concave—with faster growth when the firm is small and slower growth when the firm grows. Both the models with and without knowledge constraints are able to replicate these concave growth paths.

The difference between the models becomes evident when concentrating on small firms with very high productivity. Table 2 presents the knowledge growth paths for a simulated firm which is “born” small (low K and N), but with a very high productivity (high Z).

Panel (a) presents the growth path of the firm in a model without knowledge constraints, while Panel (b) presents the growth path in a model with knowledge constraints. While the growth path of such “tail” firms is concave without knowledge constraints, the model with knowledge constraints yields a flat growth path—Gibrat’s law holds for this type of firms in a model with knowledge constraints. Importantly, the median firm in both models is subject to a concave growth path, so this prediction only relates to the special case of high-growth-potential firms (e.g., Google’s first years).

This implication of the model is testable. The main challenge is availability of data on number of employees for private firms. An anecdotal test of the employee growth paths of both Google and Facebook, for which employee counts are available since inception (due to their high profile) yields remarkably smooth growth paths, consistent with Gibrat’s law and the predictions of the model with knowledge constraints. This is of course merely anecdotal, and a more systematic test is pending better data.

## 6. Summary

I incorporate a micro-founded constraint on knowledge investment into a standard q-theory model of a firm with two factors of production, physical and knowledge capital. I empirically test the importance of the constraint by estimating the model using data on public R&D-performing firms. My findings indicate that while financial frictions are important in explaining firm dynamics when knowledge constraints are absent, they no longer play a significant role when the model accounts for knowledge constraints.

The model with knowledge constraints is able to reproduce several salient features of the R&D data. The marginal return to knowledge investment in the estimated model is 0.217, more than double the return to physical capital. This is because firms forgo some positive NPV projects due to capacity constraints on investment in R&D. These constraints arise because firms need to transfer firm-specific knowledge to newly-hired employees, in a

training process which depends on the existing stock of R&D researchers. The model with knowledge constraints also captures the negative relation between firm size and innovation intensity, the persistence of R&D investment, and the fact that small firms exhibit higher marginal returns to R&D than large firms. All these features of the R&D data are not matched in a model lacking knowledge constraints.

I provide a test of external validation based on the number of graduating STEM PhDs relative to states' populations, which captures different firm-level training costs between states. Observed variations in firm ratios with respect to this measure correspond well to implied elasticities derived from the estimated model.

An implication of the knowledge constraints model is that capital infusion “earmarked” for R&D will not help alleviate the binding constraint for constrained firms. This is in contrast to extant models, such as Acemoglu et al. (2016), which exogenously assumes firms are born with high returns to R&D and then randomly devolve to having low returns. Knowledge constraints instead highlight the importance of the acquisition channel, which allows firms to “acqui-hire” entire R&D teams, thus easing the constraint. My findings highlight the importance of knowledge constraints to firm growth dynamics and are able to explain a relatively large set of stylized facts within a simple, intuitive model.

## References

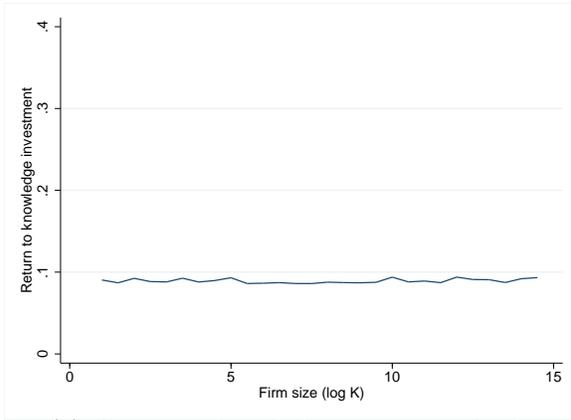
- Acemoglu, D., Akcigit, U., Bloom, N., Kerr, W. R., 2016. Innovation, Reallocation and Growth. mimeo .
- Akcigit, U., 2010. Firm size, innovation dynamics, and growth. Working paper .
- Akcigit, U., Kerr, W. R., 2010. Growth through heterogenous innovations. NBER working paper No. 16443 .
- Altinkilic, O., Hansen, R. S., 2000. Are there economies of scale in underwriting fees? Evidence of rising external financing costs. *Review of Financial Studies* 13, 191–218.
- Asquith, P., Mullins, D., 1986. Equity issues and offering dilution. *Journal of Financial Economics* 25, 61–89.
- Bradshaw, M. T., Richardson, S. S., Sloan, R. G., 2006. The relation between financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics* 42, 53–85.
- Chava, S., Roberts, M. R., 2008. How Does Financing Impact Investment ? *Journal of Finance* 63, 2085–2121.
- Cooley, T. F., Quadrini, V., 2001. Financial markets and firm dynamics. *American Economic Review* .
- Cooper, R. W., Haltiwanger, J. C., 2006. On the nature of capital adjustment costs. *Review of Economic Studies* 73, 611–633.
- DeAngelo, H., DeAngelo, L., Whited, T. M., 2011. Capital structure dynamics and transitory debt. *Journal of Financial Economics* 99, 235–261.
- Erickson, T., Whited, T. M., 2000. Measurement Error and the Relationship between Investment and  $q$ . *Journal of Political Economy* 108, 1027–1057.

- Evans, D. S., 1987. Tests of alternative theories of firm growth. *Journal of Political Economy* 95, 657–674.
- Fama, E. F., French, K. R., 1997. Industry costs of equity. *Journal of Financial Economics* 43, 153–193.
- Fazzari, S. M., Hubbard, R. G., Petersen, B. C., 1988. Financing constraints and corporate investment. *Brookings Papers on Economic Activity* 1, 141–206.
- Gomes, J. F., 2001. Financing investment. *American Economic Review* 91, 1263–1285.
- Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. *Journal of Applied Econometrics* 8, 85–118.
- Gourieroux, C. A., Monfort, A., 1996. *Simulation Based Econometric Methods*. Oxford University Press, Oxford, UK.
- Griliches, Z., 1979. Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics* 10, 92–116.
- Griliches, Z., 1998. R&D and Productivity: The Unfinished Business. In: Griliches, Z. (ed.), *R&D and Productivity: The Econometric Evidence*, University of Chicago Press, Chicago, IL, pp. 269–283.
- Hall, B. H., Hayashi, F., 1989. Research and Development As An Investment. NBER Working Paper Series 2973.
- Hall, B. H., Jaffe, A., Trajtenberg, M., 2001. The NBER patent citation data file: Lessons, insights and methodological tools. NBER working paper .
- Hall, B. H., Lerner, J., 2010. The financing of R&D and innovation. *Handbook of the economics of innovation* .

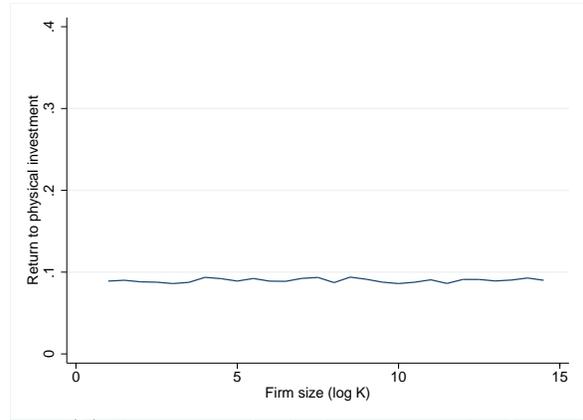
- Hall, B. H., Mairesse, J., Mohnen, P., 2010. Measuring the Returns to R&D. *Handbook of the Economics of Innovation* 2, 1033–1082.
- Hayashi, F., 1982. Tobin’s marginal  $q$  and average  $q$ : A neoclassical interpretation. *Econometrica* 50, 213–225.
- Hayashi, F., Inoue, T., 1991. The Relation Between Firm Growth and  $Q$  with Multiple Capital Goods : Theory and Evidence from Panel Data on Japanese Firms. *Econometrica* 59, 731–753.
- Hennessy, C., Levy, A., Whited, T., 2007. Testing  $Q$  theory with financing frictions. *Journal of Financial Economics* 83, 691–717.
- Hennessy, C. A., Whited, T. M., 2005. Debt dynamics. *Journal of Finance* 60, 1129–1165.
- Itenberg, O., 2014. Firm Size, Equity Financing and Innovation Activity. mimeo .
- Jagannathan, R., Matsu, D. A., Meier, I., Tarhan, V., 2016. Why do firms use high discount rates? *Journal of Financial Economics* .
- Jensen, M. C., Meckling, W. H., 1995. Specific and general knowledge, and organizational structure. *Journal of Applied Corporate Finance* .
- Jermann, U. J., 1998. Asset pricing in production economies. *Journal of Monetary Economics* 41, 257–275.
- Kalcheva, I., McLemore, P., Pant, S., 2016. Innovation: the interplay between demand-side shock and supply-side environment. Working paper .
- Krasker, W. S., 1986. Stock price movements in response to stock issues under asymmetric information. *Journal of Finance* 41, 93–105.
- Kung, H., Schmid, L., 2015. Innovation, Growth, and Asset Prices. *The Journal of Finance* 70, 1001–1037.

- Kydland, F. E., Prescott, E. C., 1982. Time to build and aggregate fluctuations. *Econometrica* 50, 1345–1370.
- Li, E. X. N., Liu, L. X. L., 2012. Intangible Assets and Cross-Sectional Stock Returns: Evidence from Structural Estimation. mimeo .
- Mincer, J., 1962. On-the-job training: costs, returns, and some implications. *Journal of Political Economy* 70, 50–79.
- Moris, F., 2005. The U.S. Research and Experimentation Tax Credit in the 1990s. National Science Foundation .
- OECD, 2012. Business enterprise R-D expenditure by industry and by type of cost. OECD Science, Technology and R&D Statistics .
- Penrose, E., 1959. The theory of the growth of the firm. John Wiley & Sons, New-York, NY.
- Phillips, G., Zhdanov, A., 2011. R&D and the Market for Acquisitions. mimeo .
- Prescott, E. C., Visscher, M., 1980. Organization Capital. *Journal of Political Economy* .
- Rauh, J. D., 2006. Investment and Financing Constraints : Evidence from the Funding of Corporate Pension Plans. *Journal of Finance* 61, 33–71.
- Schumpeter, J. A., 1942. *Capitalism, Socialism and Democracy*. Routledge, London.
- Sevilir, M., Tian, X., 2011. Acquiring Innovation. mimeo .
- Uzawa, H., 1965. Optimum technical change in an aggregative model of economic growth. *International economic review* 6, 18–31.
- Warusawitharana, M., 2014. Research and Development, Profits and Firm Value : A Structural Estimation. mimeo .

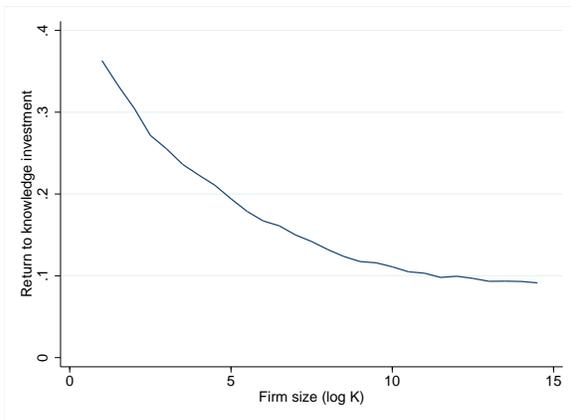
- Weiss, L. A., 1990. Bankruptcy resolution: Direct costs and violation of priority of claims. *Journal of Financial Economics* .
- Whited, T. M., 1992. Debt, liquidity constraints, and corporate investment: evidence from panel data. *Journal of Finance* .
- Wildasin, D. E., 1984. The q theory of investment with many capital goods. *American Economic Review* 74, 203–210.



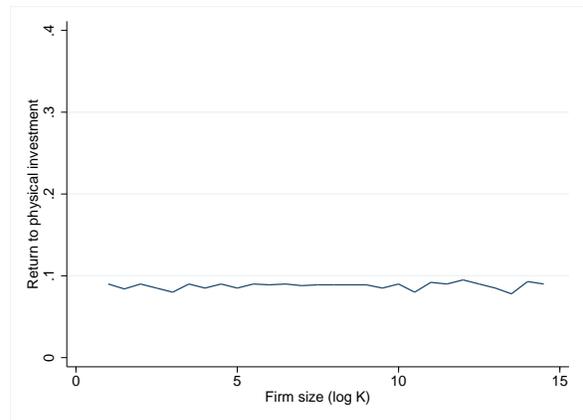
(a) Return to knowledge - unconstrained



(b) Return to physical - unconstrained

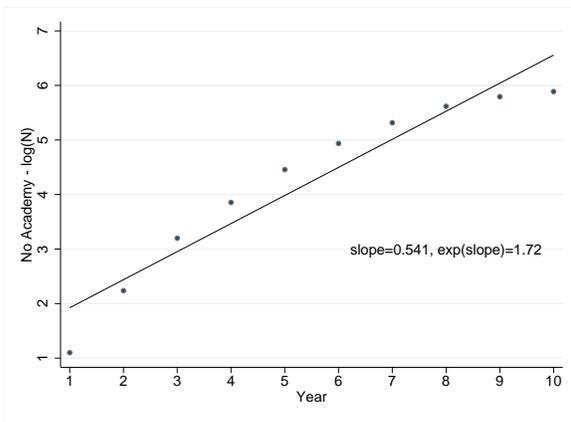


(c) Return to knowledge - constrained

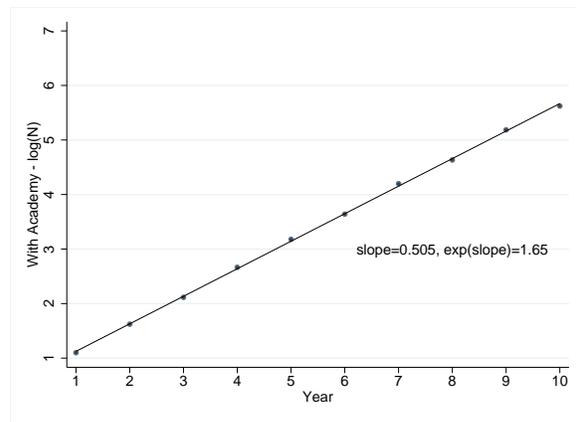


(d) Return to physical - constrained

Fig. 1. Returns to investment. This figure presents the average (across firms) of the marginal return to investment in knowledge capital and physical capital in the simulated panel, by firm size. Panels (a) and (b) present the respective returns derived from a model without a knowledge constraint. Panels (c) and (d) present the respective returns derived from a model with a knowledge constraint. Returns are derived directly from the value functions of the estimated models.



(a) Simulated growth path - no constraint



(b) Simulated growth path - with constraint

Fig. 2. Simulated growth paths. This figure presents the simulated growth paths of a small firm (small  $K$  and  $N$ ) with high productivity (high  $Z$ ). Panel (a) presents the simulated growth path in the model without a knowledge constraint. Panel (b) presents the simulated growth path when the model includes a knowledge constraint.

Table 1  
Variable definitions

This table describes the definitions of the variables used in the paper. Item names are Compustat mnemonics: ppegt - plant property and equipment gross total; sale - sales; cogs - cost of goods sold (including payment to production workers); xsga - selling, general and administrative (includes R&D expenditure); xrd - R&D expenditure; capx - capital expenditure; sppe - sale of plant property and equipment; acq - acquisitions; dvc,dvp - dividends; prstkc - stock repurchase; xint - interest payments; at - assets total; seq - shareholders equity; txditc - tax credits; pstk - preferred stock; prcc.f - closing price at end of year; csho - stocks outstanding. The coefficients  $\psi_{i,t}^N$  and  $\psi_{i,t}^K$  are acquisition imputation coefficients which are firm-year specific, such that  $\psi_{i,t}^N + \psi_{i,t}^K = 1$ , and are derived from the intensities of physical and knowledge investments pre-acquisition.

ID	Description	Data (Compustat mnemonics)
K <sup>1</sup>	Physical capital	ppegt
N	Knowledge capital	/
Z	Productivity	/
CF	Cashflow from operations	sale - cogs - (xsga - xrd)
IK	Physical investment (net of sale)	capx - sppe + $\psi_{i,t}^K$ acq
IN	Knowledge investment (R&D)	xrd + $\psi_{i,t}^N$ acq
D <sup>2</sup>	Distributions	(dvc + dvp + prstkc - sstk) + (xint - $\Delta$ (at - (seq+txditc-pstk)))
V <sup>1</sup>	Total firm value	prcc.f.csho + (at - (seq+txditc-pstk))

<sup>1</sup> Lagged one year, to be beginning of period value.

<sup>2</sup>  $\Delta$  denotes current minus lagged value.

Table 2  
Moments

This table presents data moments (Data), simulated moments from a model without a knowledge constraint (No Acad), and simulated moments from a model with a knowledge constraint (Acad), along with t-statistics on the difference between the data moments and the simulated moments. For variable definitions, see Table 1. “Std.” is the standard deviation of the relevant variable. Growth rates are in log terms. Decrease in R&D and physical intensity are the slope coefficients from regressions of log R&D to capital intensity and log physical investment to capital intensity on firm size (log K), respectively.

Description	Data	No Acad	t-stat	Acad	t-stat
Mean physical capital (log K)	5.116	5.562	1.693*	5.610	1.798*
Std. log K	2.031	2.212	1.478	2.187	1.276
Mean growth rate log K	0.067	0.064	-0.143	0.063	-0.185
Std. growth rate log K	0.204	0.181	-0.715	0.154	-1.573
Mean <sup>1</sup> value-to-capital ratio (V/K)	4.020	1.471	0.816	1.275	-1.173
Std. growth rate V/K	0.425	0.415	-0.217	0.395	-0.641
Mean cashflow-to-capital ratio (CF/K)	0.550	0.614	0.752	0.565	0.179
Persistence of CF/K	0.876	0.814	-0.603	0.795	-0.780
Mean distributions-to-value ratio (D/V)	0.033	0.050	1.541	0.035	0.186
Std. of D/V	0.105	0.058	-2.126**	0.049	-2.528**
Mean <sup>1</sup> R&D-to-capital ratio (IN/K)	0.144	0.121	-1.026	0.133	-0.469
Persistence of log IN/K	0.931	0.841	-2.487**	0.937	0.171
Decrease in R&D intensity (log IN/K on log K)	-0.167	-0.025	3.617***	-0.170	-0.068
Std. growth rate of R&D	0.491	0.455	-0.407	0.424	-0.765
Std. growth rate of R&D intensity (IN/K)	0.536	0.464	-0.752	0.440	-1.001
Mean <sup>1</sup> phys. investment-to-capital ratio (IK/K)	0.118	0.106	-1.267	0.128	0.837
Persistence of log IK/K	0.583	0.856	3.344***	0.820	2.908***
Decrease in phys. investment intensity (log IK/K on log K)	-0.029	-0.043	-0.442	-0.048	-0.615
Std. growth rate of phys. investment	0.699	0.780	1.243	0.668	-0.481
Std. growth rate of phys. investment intensity (IK/K)	0.766	0.801	0.468	0.694	-0.973

\*\*\*, \*\*, \* Difference significant at the 1%, 5% and 10% level, respectively.

<sup>1</sup> More precisely,  $\exp(\text{mean}(\log(X)))$ , which better corresponds to  $\text{median}(X)$  for log-normal  $X$ .

Table 3  
Estimated parameters

This table presents the estimated parameter values for the models without a knowledge constraint (No Acad) and with a knowledge constraint (Acad), along with their standard errors (s.e.). Also reported are the over-identification J-statistics for both models, along with the p-values on rejecting each model in brackets.

Param	Description	No Acad	s.e.	Acad	s.e.
$\theta$	Production return to scale	0.784	0.137	0.879	0.154
$\omega$	Elasticity of physical capital	0.678	0.191	0.676	0.197
$\rho_Z$	Productivity persistence	0.949	0.069	0.928	0.065
$\mu_Z$	Average TFP	0.569	0.244	0.052	0.251
$\sigma_Z$	Std. dev. of shock to productivity	0.327	0.164	0.376	0.172
$\delta_K$	Physical capital depreciation	0.113	0.015	0.110	0.014
$\lambda_K$	Physical capital adjustment	0.668	0.226	0.606	0.259
$\delta_N$	Knowledge capital depreciation	0.248	0.083	0.244	0.091
$\lambda_N$	Knowledge capital adjustment	0.288	0.110	0.479	0.099
$\alpha$	Knowledge constraint parameter	-	-	2.529	0.941
$\gamma$	Financial friction parameter	0.760	0.127	0.073	0.238
$r$	Expected rate of return	0.093	0.021	0.093	0.020
Model rejection J-test		15.422	(0.073)	10.317	(0.247)

Table 4  
Labor training tests

This table presents the results of several regressions of firm variables on a proxy for labor skill at the state where the firm is headquartered. For firm variable definitions, see Table 1. The proxy is constructed from the number of PhD graduates in the state relative to the state's population, and is standardized. Also reported are the t-stat and adjusted-R<sup>2</sup> of each regression. The number of observation in the regressions is 34,792. All regressions include year and industry fixed effects. The last column reports the elasticity of the variables to an increase in  $\alpha$  (easing the knowledge constraint) within the estimated model with knowledge constraints.

Variable	Coeff	t-stat	Adj-R <sup>2</sup>	Model elasticity
Growth in physical capital (log K)	0.00125	0.87	0.0299	0.420
Growth in firm value (log V)	0.00526	2.19	0.1022	1.605
Growth in cashflow (log CF)	0.00668	1.71	0.0125	1.281
Growth in R&D (log IN)	0.00429	1.90	0.0089	1.719
Growth in phys. investment (log IK)	0.00476	1.09	0.0231	1.351
Value to capital ratio (log V/K)	0.02706	5.62	0.2712	0.426
Cashflow to capital ratio (log CF/K)	0.02679	4.87	0.1642	0.448
R&D to capital ratio (log IN/K)	0.07481	11.58	0.4104	1.845
Investment to capital ratio (log IK/K)	0.01832	3.62	0.0869	0.477
Cashflow to value ratio (log CF/V)	-0.00304	-0.71	0.0580	-0.503
R&D to value ratio (log IN/V)	0.04775	9.35	0.2632	0.917
Investment to value ratio (log IK/V)	-0.00823	-1.59	0.2079	-0.108

## Appendix A. Appendix

### A.1. Data distributions

It is well known that the firm size distribution is exponentially distributed. As a consequence, so are all other firm variables, which typically scale with firm size (such as investments, cash flows, etc). For that reason, it is common in the structural modeling literature to consider moments of the intensity distributions for these variables—dividing all variables by firm size before calculating moments of the their distributions.

The distributions of firm intensity variables (such as physical investment intensity -  $IK/K$ ) is however exponentially distributed as well. This is easy to see from Fig. A.1, which plots the distribution of the *log* of physical and R&D intensity in the data. Both log intensities are observationally normally distributed, implying the intensities themselves are close to log-normally distributed.

The two primary distribution moments used in structural estimation are the mean and standard deviation of the distribution. For exponentially distributed variables, estimators for both of these moments are strongly affected by tail observations. The estimators are hence noisy, making the standard errors on the estimators fairly large. The large standard errors are especially problematic in the context of structural modeling, because it is established practice to use the inverse-variance-covariance matrix of the data moments as the weighting matrix of the SMM procedure (as described in Section 4.1).

The result of using estimators with large standard errors is that the weight assigned by the estimation procedure to these moments is then much lower than the weight assigned to better measured estimators, such as regression coefficients. The under-weighting of these main moments of the data distributions leads to less efficient inference. I therefore take a log transform of all positive exponentially distributed intensities, and fit moments of the log intensity distributions rather than moments of the intensity distributions.

## *A.2. Firm dynamism*

TODO - discussion of the dynamism (rates of exit and entry) exhibited by the data, and presentation of Appendix Fig. A.2.

## *A.3. Moment sensitivities*

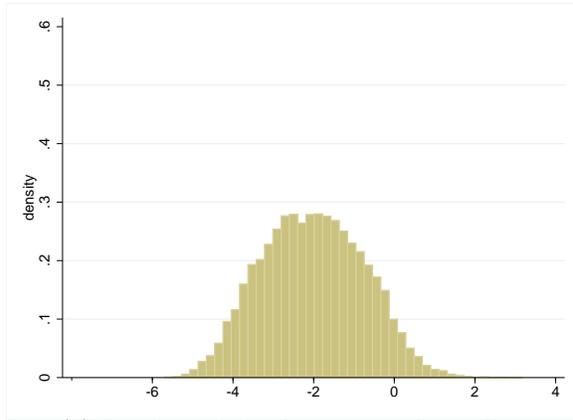
TODO - discussion of the sensitivity of the used moments to the parameters of the model. Include the table of elasticity of moments to parameters, and the graph of local identification of the  $J$ -statistic.

## *A.4. Counterfactual physical constraints*

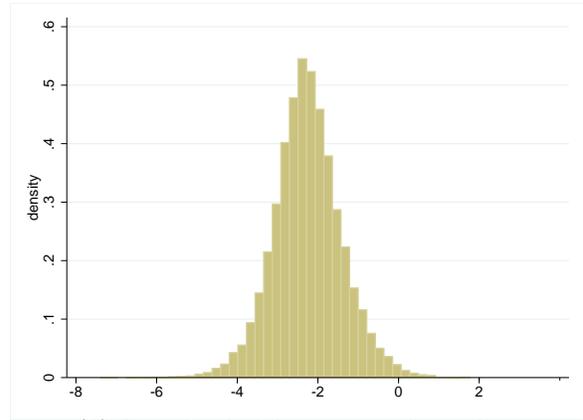
TODO - discussion of the counterfactual in which the “firm academy” constraint is imposed on physical capital accumulation, or on both knowledge capital and physical capital accumulation. Present Appendix Table A.1.

## *A.5. Acquisition activity*

TODO - Discuss estimation with different approaches to acquisition activity by firms, and imputation of acquisitions into investments in physical and knowledge capital using the goodwill proxy.

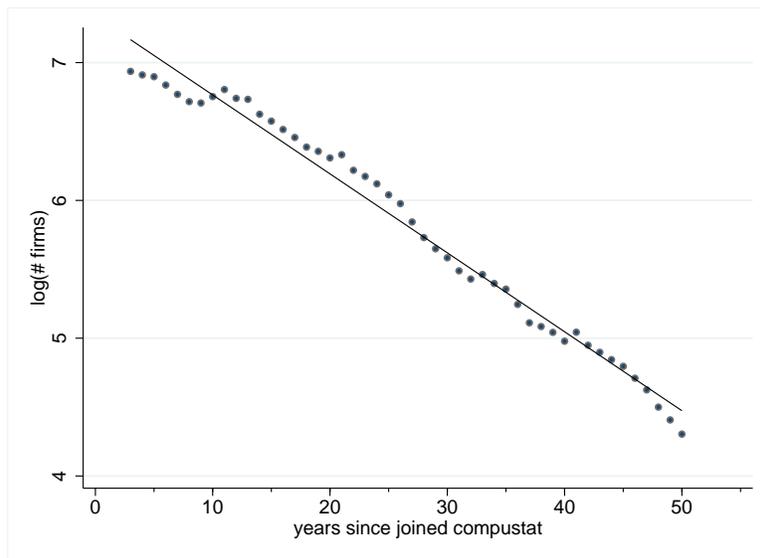


(a) Log knowledge investment intensity



(b) Log physical investment intensity

Fig. A.1. Distributions of log investment intensities. Panels (a) and (b) of this figure present the histograms of knowledge and physical investment intensities in the data, respectively. The log of knowledge investment intensity is  $\log(IN/K) = in - k$ , with  $IN$  and  $K$  defined in Table 1, and lowercase letter denoting logs. Log physical investment intensity is similarly defined as  $ik - k$ .



(a) Firm age decay

Fig. A.2. Decay of number of firms with age. This figure presents the log of the number of firms at each age (defined as the time since the firm first appeared in Compustat) at the end of 2014, for the R&D performing firms sample. The regression line is also plotted, and its slope is -0.064.

Table A.1  
 Estimated parameters - physical vs. knowledge constraints

This table presents the estimated parameter values for the models with only physical constraints (Phys. Acad), and with both physical and knowledge constraints (Both Acad), along with their standard errors (s.e.). Also reported are the over-identification J-statistics for both models, along with the p-values on rejecting each model in brackets.

Param	Description	Phys. Acad	s.e.	Both Acad	s.e.
$\theta$	Production return to scale	0.764	0.139	0.868	0.155
$\omega$	Elasticity of physical capital	0.675	0.185	0.676	0.194
$\rho_Z$	Productivity persistence	0.948	0.069	0.923	0.068
$\mu_Z$	Average TFP	0.586	0.224	0.049	0.260
$\sigma_Z$	Std. dev. of shock to productivity	0.346	0.159	0.375	0.170
$\delta_K$	Physical capital depreciation	0.112	0.014	0.110	0.014
$\lambda_K$	Physical capital adjustment	0.724	0.218	0.666	0.273
$\delta_N$	Knowledge capital depreciation	0.246	0.086	0.221	0.097
$\lambda_N$	Knowledge capital adjustment	0.303	0.115	0.471	0.102
$\alpha_N$	Knowledge constraint parameter	-	-	2.607	0.975
$\alpha_K$	Physical constraint parameter	22.911	6.303	27.006	7.858
$\gamma$	Financial friction parameter	0.748	0.138	0.077	0.247
$r$	Expected rate of return	0.093	0.022	0.092	0.019
Model rejection J-test		15.436	(0.052)	10.298	(0.172)