R&D Return Dispersion And Growth *

Nils H. Lehr†
Boston University

September 13, 2022
Click here more the latest version.

Abstract

This paper documents large and persistent differences in R&D returns across listed US firms, with firms at the 75th percentile having two times the median return. This dispersion is surprising as workhorse endogenous growth models predict that R&D resources flow from low to high return firms until return equalization. I investigate frictions as a potential mechanism and find mixed results. On the one hand, I document that high R&D return firms are not credit constrained nor have a higher return on capital, suggesting that financial frictions are not a key driver. On the other hand, I document that high R&D return firms face a more inelastic supply of inventors, suggesting that return differences might reflect labor-market power. Calibrating a Schumpeterian growth model to match key data moments, I find that inventor market power differences can explain 1/3 of the measured R&D return dispersion and reduce the annual growth rate by 0.06 p.p., thereby reducing welfare by 2.1%.

*I especially thank Stephen J. Terry, Pascual Restrepo, Tarek A. Hassan, and David Lagakos for their continued guidance, feedback, and encouragement. I benefited from the advice and comments of many including Stefania Garetto, Adam Guren, Maarten De Ridder, Diego Restuccia, Sara Moreira, Dimitris Papanikolaou, Stefanie Stantcheva, Titan Alon, and the participants of the Macro Lunch Seminar and Green Line Macro Meeting.

†Email: nilslehr@bu.edu Mailing Address: Boston University, Dept. of Economics, 270 Bay State Rd., Room 515, Boston, MA 02215.

1 Introduction

Innovation created by research and development (R&D) is considered an important driver of modern economic growth. In the past twenty years, the US has experienced declining economic growth while R&D expenditure has remained stable. This finding has led to a renewed interest in the determinants of aggregate R&D productivity, broadly defined as the gap between aggregate R&D expenditure and economic growth.

In the US, more than 70% of R&D is performed by businesses, which arguably conduct it to profit from the resulting innovation and earn a return on their investment.[1] The return on R&D, therefore, may determine the allocation of R&D activity across firms, which a key determinant of aggregate R&D productivity and, thus, economic growth.

In this paper, I document that R&D returns, which I measure as the ratio of patent valuations to R&D expenditure, are highly dispersed across listed firms in the US. A firm at the 75th percentile earns almost double the median return. This dispersion is surprising as workhorse endogenous growth models predict that R&D resources flow from low to high return firms until returns are equalized, thereby maximizing the total value created.[2] Return dispersion then suggests that the economy is not allocating its R&D resources most efficiently from a private perspective, which, as long as firms' incentives are sufficiently aligned with economic growth, also implies a failure to maximize the economy's growth potential.

I investigate frictions as a potential mechanism preventing R&D resource reallocation and find mixed results. On the one hand, standard financial frictions do not appear to drive the measured dispersion. R&D returns are uncorrelated with the return on capital, a proxy for investment frictions, and correlation patterns with other proxies of financial frictions do not support the conclusion that financial constraints drive R&D return dispersion.

On the other hand, I document that distortions in the market for inventors may contribute to R&D return dispersion. Under this mechanism, firms facing more inelastic inventor labor supply suppress their hiring to keep wages low and, as a by-product, achieve high returns. I estimate that differences in the labor supply elasticity can account for 64% of the return difference between firms with above and below median R&D return. Thus, R&D return dispersion might not be a failure from firms' perspective, but a signal that market power heterogeneity affects the allocation of scarce R&D resources.

[1] According to the NSF National Patterns, the average share of R&D performed by business between 1975 and 2020 is 71%. The average share of R&D funded is slightly lower at 59% and has been steadily increasing.

[2] See for example Acemoglu et al. (2018); Akcigit and Kerr (2018); de Ridder (2021); Aghion et al. (2022b).

I estimate the impact of these forces on R&D return dispersion and economic growth in a calibrated, quantitative Schumpeterian growth model. In the model, firms with heterogeneous R&D productivity hire inventors subject to adjustment costs and monopsony power. The latter is modeled so that the labor supply elasticity is decreasing in the firm's inventor demand, as in a popular recent approach in Card et al. (2018).

I calibrate the model using a combination of parameters chosen from either the literature or moment matching. The model can account for 1/3 of the R&D return dispersion and predicts an increase in the annual growth-rate of 0.06 p.p. in the absence of inventor market distortions, a consumption equivalent welfare improvement of 2.1%. For comparison, Hsieh and Klenow (2009) estimate welfare cost of capital misallocation around 30%-40% for the US, while Berger et al. (2022) estimate the cost of labor monopsony in the production sector at 7.6%. More generally, Lucas (2003) and Arkolakis et al. (2012) estimate that the welfare cost of business cycles and trade autarky for the US are around 1%. The documented dispersion in R&D returns, thus, informs modeling of the R&D sector and may be quantitatively important for economic growth.

Literature. This paper contributes to four strands of the literature. First, I provide new findings potentially linked to the drivers of aggregate R&D resources allocation and productivity. While existing literature focuses on the misalignment of private and public incentives in R&D, I show that private frictions may matter as well.[3] In particular, I document large dispersion in R&D returns, which workhorse models interpret as evidence in favor of significant frictions in the market for R&D resources.[4] My estimates suggests that labor market frictions are quantitatively important, while financial frictions may be less so.

Second, the documented R&D return dispersion speaks to the literature on factor misallocation.[5] Focusing on R&D instead of static production factors such as capital and labor introduces a dynamic component linking return heterogeneity to the growth rate instead of static production efficiency. Furthermore, I highlight the importance of labor market frictions in the market for inventors as an important driving factor. The existing literature primarily

[3]While early models focused on aggregate under-provision of R&D, recent advances emphasize sub-optimal allocation of R&D resources across firms due to differences in their ability to exploiting innovation. See e.g. Romer (1990); Aghion and Howitt (1992); Acemoglu and Cao (2015); Acemoglu et al. (2018); Akcigit and Kerr (2018); Peters (2020); de Ridder (2021); Aghion et al. (2022b); Terry (2022).

[4]Akcigit et al. (2022) also consider private frictions when investigating the optimal design of R&D tax credits. They calibrate frictions in a reduced form using the ratio of sales changes to R&D expenditure.

[5]See Restuccia and Rogerson (2008); Hsieh and Klenow (2009); Midrigan and Xu (2014); David et al. (2016); David and Venkateswaran (2019); David et al. (2021) for more recent advances.

studies investment frictions including adjustment cost and financial frictions, which I show might be of lesser importance in the context of R&D investments.

Third, my paper is closely related to the growing literature on labor market frictions and resource allocation.[6] In line with recent findings, I provide evidence that heterogeneity in the labor supply elasticity across firms is important for factor allocation, however, in a new context: the market for inventors. Thus, these forces harm innovation and economic growth as well as static production efficiency. Furthermore, I build on the growing literature linking heterogeneity in the labor supply elasticity to firm size, and quantify the importance of this link for economic growth (Seegmiller, 2021; Berger et al., 2022). My estimates suggest that these labor market frictions are quantitatively important for economic growth.

Fourth, my measure of R&D returns builds on recent advances in the literature on measuring innovation. I quantify the private value created from R&D using the patent valuation measure developed in Kogan et al. (2017), who estimate it using an event study estimation approach. Focusing on the value created for the firm is crucial as it directly connects to their incentives. I contribute to this literature by carefully measuring firm-level R&D returns, documenting their empirical dispersion, and providing evidence on their underlying drivers. My findings highlight the importance of understanding the empirical distribution of returns on R&D for the study of economic growth and development of innovation policy.

Structure. Section 2 introduces the data. Section 3 documents R&D return dispersion and Section 4 investigates potential drivers thereof. Section 5 estimates the quantitative importance of inventor labor market frictions for R&D return dispersion and economic growth. Section 6 concludes.

[6]See (Card et al., 2018; Azar et al., 2019; Kline et al., 2019; Shi, 2020; Seegmiller, 2021; Berger et al., 2022; Yeh et al., 2022)

2 Data

My data combines information on the financial performance and innovations of US listed firms. In this section I briefly describe its construction with further details in Appendix A.

I obtain balance sheet and income statement data for US listed firms from WRDS Compustat. These data are taken directly from mandatory filings by the company and harmonized by Compustat. The data reaches back to 1959 and its availability is tied to the listing status of the company. For my purposes, the most important variables are sales (`sale`), employment (`emp`), capital stock (`ppent`), and R&D expenditure (`xrd`). Following Peters and Taylor (2017), I measure the capital stock at the beginning of period.

I construct firm-level measures of innovation including total new patents and their valuation using the data in Kogan et al. (2017). Their data links granted patents to US listed firms and estimates their value based on the stock market response to their announcement by the US Patent and Trademark Office (USPTO). I record patents in their application year and aggregate to the firm-year level.

Patents are arguably the most direct measure of R&D output available to researchers. A patent captures an invention that the issuing patent office, here the USPTO, deemed new and useful, and grants the owner exclusive rights to the use of the invention described therein. These rights give firms strong incentives to patent inventions, making the value of newly granted patents a direct measure of their innovation output.[7] Patent valuations, in turn, capture the private value of an invention, which is directly linked to firms' incentives to innovate. In contrast, other patent-based measures of innovation such as (citation-adjusted) patent counts capture the quantity of innovation, but not necessarily its value to the firm.[8]

I measure inventor employment, which is important for my analysis on labor market imperfections, at the firm-level using patent records. I first identify and link inventors across patents using the disambiguation provided by the USPTO. I then assign them to the firm based on whether they are listed on a firm's newly-granted patent within the relevant 5-year window. I assign the firm a full time equivalent share of the inventor using the ratio of patents linked to the firm and inventor divided by all patents linked to the inventor within the relevant time window. I aggregate to the firm-level by summing over all inventors.

[7] Not all inventions are patented (Cohen and Klepper, 1996). Thus, patents remain an imperfect measure.

[8] These concepts can diverge e.g. due to externalities or because some firms are better equipped than other to take advantage of an invention. See e.g. Lerner (1995); Bloom et al. (2013); Kogan et al. (2017); Akcigit and Kerr (2018); de Ridder (2021); Kelly et al. (2021); Aghion et al. (2022b)

The combined data thus has information on firms' financial decisions, their inventions, and their inventor employment. I restrict my sample to the 1975-2014 period and drop firms with consistently low R&D expenditure (less than 2.5m 2012 USD per year), low patenting (less than 2.5 patents per year) or less than 5 active years in the sample. The final sample covers more than 80% of R&D expenditure in Compustat and patent valuations in Kogan et al. (2017) for the 1975-2014 period as well as $\approx 40\%$ of the R&D recorded in BEA accounts.

3 Documenting Return on R&D Dispersion

This section introduces the measurement of R&D returns and documents their dispersion.

3.1 Measurement

The return on R&D is defined the ratio of the value created from R&D divided by its cost. Conceptually, we can attribute variation in this measure to two potential sources: variation in the expected returns at the time of investment, and variation around this value once the associated projects are completed and their value is revealed. While the former is informative about the R&D decision-making process, the latter primarily speaks to the extend of uncertainty in innovation. In this paper, I will focus on the R&D decision making process and, hence, construct measures of expected R&D returns. I will measure costs from R&D expenditure, and, as discussed above, R&D output using patent valuations.

I measure the Expected Return on R&D for firm i in year t as the ratio of patent valuations to previous year's R&D expenditure at the 5-year horizon:

$$\texttt{Expected Return on R\&D}_{it} \equiv \frac{\sum_{s=0}^{4} \texttt{Patent Valuations}_{it+s}}{\sum_{s=0}^{4} \texttt{R\&D Expenditure}_{it-1+s}}. \tag{1}$$

I drop observations based on less than 50 patent valuations. The median (average) return has around 160 (520) underlying patent valuations. Focusing on an extended horizon with many underlying patents allows me to measure ex-ante expectations following standard law of large numbers arguments.[9] My final sample has around 12,000 returns from 900 firms.

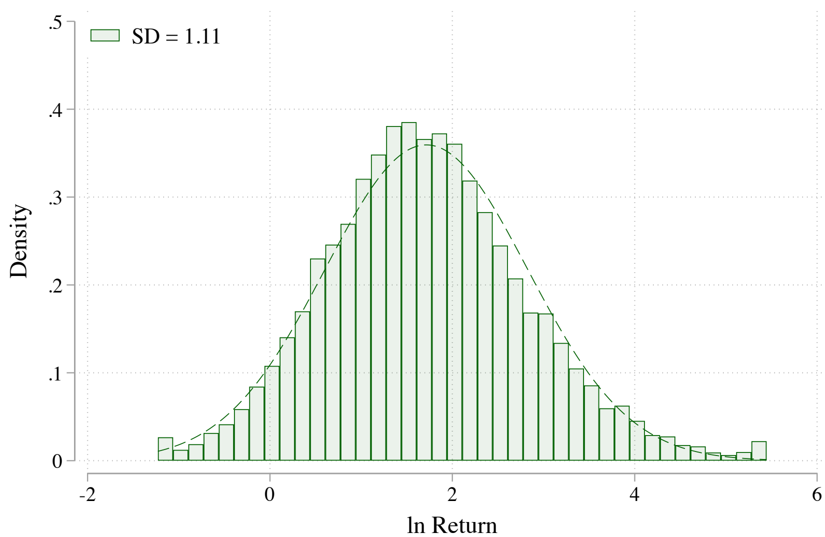
[9]Note, however, that patent valuations are tied to USPTO patent grant announcements, which occur on a weekly basis. The minimum different announcement weeks for a return in my sample is 28 with a median (average) of 117 (159).

3.2 R&D Return Dispersion

The histogram in Figure 1 shows that expected returns on R&D are widely dispersed. A firm at the 75th percentile of the distribution has a 1.9 times larger expected return compared to the median. The standard deviation of the expected return to R&D is 1.1.

The identity of firms across the distribution is noteworthy as well. For example, many successful tech-firms including Apple, Alphabet, and Qualcomm consistently earn high returns, while legacy industry giants such as General Motors and Ford Motors consistently earn below median returns. Appendix E reports firms with the highest and lowest average returns in my sample.

Figure 1: Returns on R&D are Dispersed



Notes: This figure plots a histogram of the demeaned natural logarithm of expected R&D returns. SD refers to the standard deviation. The dashed line plots the density function of a normal distribution with the same mean and standard deviation. R&D returns are measured as the 5-year total patent valuation divided by 5-year R&D expenditures lagged by one year. See Section 2 and Appendix A for data detail.

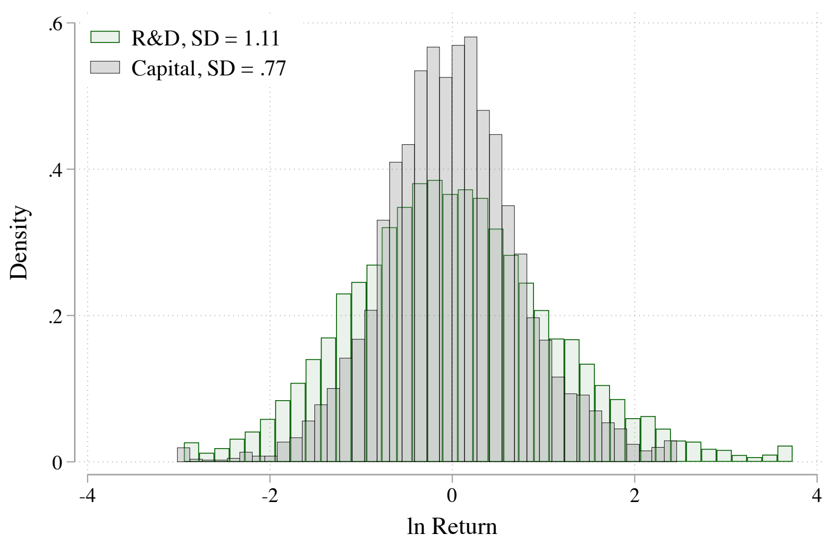
I benchmark the measured dispersion using the return on capital, which, following David et al. (2021), is defined as the ratio of sales to beginning of period capital stock. As for the return on R&D, I construct the measure at the 5-year level:

$$\texttt{Return on Capital}_{it} \equiv \frac{\sum_{s=0}^{4} \texttt{Sales}_{it+s}}{\sum_{s=0}^{4} \texttt{Capital}_{it+s}}. \tag{2}$$

The return on capital is an interesting benchmark as a large literature argues that its dispersion is an indicator for misallocation of capital with potentially large cost in terms of production efficiency and welfare.[10] In theory, it should be equalized across firms as capital flows from low to high return firms to maximize overall returns. Dispersion is then a sign of inefficient investment allocation.

Comparing across return measures, I conclude that measured dispersion in the return on R&D is large. As shown in the histogram in Figure 2, it is about 40% larger than the dispersion in the return on capital and this difference is highly statistically significant.[11] Similarly, the gap between a firm at the 75th percentile and the median is c. 20% larger for the return on R&D.

Figure 2: Comparing Dispersion in the Return on R&D and Capital



Notes: This figure plots a histogram of the demeaned natural logarithm of the expected capital and R&D returns. SD refers to the standard deviation. R&D returns are measured as the 5-year total patent valuation divided by 5-year R&D expenditures lagged by one year. Capital returns are defined as 5-year sales divided by 5-year beginning of period capital stock. See Section 2 and Appendix A for data details.

3.3 Robustness

Throughout numerous robustness exercises I find that R&D returns continue to be highly dispersed and significantly more so than the return on capital. In the following I will discuss the main exercises with additional material and details in Appendix B.1.

[10] See e.g. Restuccia and Rogerson (2008); Hsieh and Klenow (2009, 2014); David et al. (2016); David and Venkateswaran (2019).

[11] The standard errors around both estimates smaller than 0.02.

Return on R&D Specification. I have three main choices in the construction of the return on R&D: the time-window in consideration, the lag between R&D expenditure and patent valuations, and the minimum number of patents required. For my baseline definition in equation (1), I chose a window of 5 years, a lag of 1 year, and a minimum of 50 patents. Table 1 confirms that neither choice is driving my results. Extending the time-window increases measured dispersion for the return on R&D, but not for alternative return measures, such that the difference is even more pronounced at longer time-windows. Similarly, extending the lag between patent valuations and R&D expenditure increases the dispersion in the measured Return on R&D. Finally, requiring at least 200 patents reduces the dispersion in the return on R&D by about 8%, however, the sample selection reduces the dispersion in the return on capital even faster such that the relative gap increases.[12]

Table 1: Return Dispersion Across Specifications

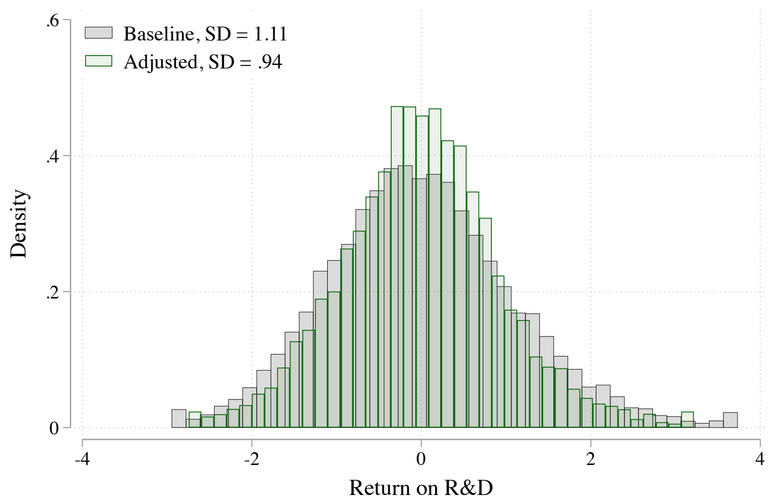
Specificaton	Return on R&D	Return on Capital	
	SD	SD	Δ%
A. Time-window			
5-year	1.11	0.77	43.4%
10-year	1.14	0.75	51.6%
20-year	1.16	0.72	61.6%
B. Minimum patents			
50 patents	1.11	0.77	43.4%
100 patents	1.04	0.71	45.3%
200 patents	1.02	0.68	50.6%
C. Realization lag			
1-year	1.11	0.77	43.4%
3-year	1.20	0.79	51.6%
5-year	1.28	0.79	62.3%

Note: Baseline specification is a horizon of 5 years with at least 50 patents and a 1-year realization lag. Dispersion calculated for sample without missing observations across return measures. Column headers SD report standard deviations of return measure. Column (3) reports the difference of Return on R&D and Capital dispersion relative to Return on Capital dispersion. Returns are measured in logs.

[12] In Appendix B.1 I allow the benefits of R&D to be distributed across time to account for the probabilistic nature of innovation. The measured dispersion is at best marginally lower under this alternative assumption.

Measurement Error. Dispersion in R&D returns could be due to measurement error arising, e.g., from the expectation-realization gap, patent valuation estimation, or misreporting of R&D expenditure.[13] I investigate this concern in two separate exercises. First, I exploit the strong auto-covariance of R&D returns to estimate the contribution of classical measurement error with a structural GMM estimator. Second, I use a bootstrapping approach to gauge the extend to which the uncertainty in patent valuation could lead to measurement error. Both exercises suggest that measurement error contributes less than 5% of the variation in the R&D returns. I thus do not find evidence that measurement error significantly contributes to R&D return dispersion. See Appendix B.2 for details.

Figure 3: Return on R&D Dispersion Before and After Adjustment



Notes: This figure plots a histogram of the demeaned natural logarithm of unadjusted and adjusted Expected Returns to R&D. Adjustments include (1) winsorizing patent valuations, (2) knowledge capital, (3) NAICS3×Year effects, (4) amenities, and (5) acquisitions. See Section 3 and B.1 for details. SD refers to the standard deviation.

Additional robustness. Appendix B.1 reports additional robustness exercises supporting the conclusion that R&D returns are highly dispersed. As shown in Figure 3, 85% of the variation remains unexplained even after a range of adjustments including (1) winsorizing patent valuations, (2) adjusted input measure, (3) NAICS3×Year effects, (4) amenities, and (5) acquisitions.[14] Most of this difference is due to industry×year effects.

[13] Note that R&D expenditure is expensed in US GAAP accounting, giving firms an incentive to fully report R&D expenditure to reduce their tax liability. Terry et al. (2022) argue that managers still might misreport when attempting to hit short-run earnings targets or smooth earnings. See also Dukes et al. (1980); Baber et al. (1991); Lev et al. (2005); Chen et al. (2021); Terry (2022).

[14] I implement these adjustments as follows: Firstly, I construct a raw return measure using winsorized patent valuations minus a share of acquisitions for R&D output and the knowledge capital measure from

4 Drivers of Return on R&D Dispersion

In this section I investigate the economic drivers of dispersion in the Return on R&D. I begin by highlighting that workhorse economic growth models predict return equalization as firms equalize marginal cost to marginal benefit. Introducing frictions allows us to reconcile theory and data. Motivated by this insight, I investigate the role of two frictions. Firstly, I investigate investment frictions and find, perhaps surprisingly, no role for them empirically. Secondly, I investigate labor market frictions due to firms' potential monopsony power in hiring inventors and find strong empirical support. In line with the theory, I find that firms with more apparent pricing power in the market for inventors have higher returns on R&D. The results thus suggests that labor market frictions, but not investment frictions, seem more important for the measured dispersion.

4.1 The Basic R&D Investment Model

Dispersion in the expected return on R&D is at odds with workhorse endogenous growth models (Romer, 1990; Aghion and Howitt, 1992; Acemoglu and Cao, 2015). In these models, firms equalize the marginal expected return on R&D when choosing their level of R&D investment by setting marginal benefit equal to marginal cost. Under standard assumptions, the marginal expected return is proportional to the total or average return up to a constant factor and, thus, the expected return on R&D is equalized across all firms.

Formally, let ℓ be the input to R&D, which could be a composite good, with associated cost function $C(\ell)$ and production function $F(\ell)$. The firm values innovation output at expected value V and solves a static maximization problem:

$$\ell^* \equiv \arg\max_{\ell} \left\{ F(\ell)V - C(\ell) \right\}. \tag{3}$$

The first order conditions of this problem can be rearranged to express the expected return on R&D, which is defined as the benefits of R&D divided by the cost, as a function of the scale elasticities of the production and cost function:

$$\text{Expected Return on R\&D} \equiv \frac{F(\ell^*)V}{C(\ell^*)} = \frac{\varepsilon_C}{\varepsilon_F}, \tag{4}$$

Ewens et al. (2020) as input measure. I determine the share of acquisitions (`aqc`) that I subtract from patent valuation in a separate step via regression, which I detail in Appendix B.1. I then residualize this return measure further with respect to NAICS3×Year effects and measures of amenities including the average temperature, housing price, and income level in the location of inventors to arrive a the final return on R&D.

where $\varepsilon_X \equiv \left.\frac{\partial \ln X(\ell)}{\partial \ln \ell}\right|_{\ell=\ell^*}$ for $X \in \{C, F\}$ is the scale elasticity of the production and cost function around the optimal research input.

The literature typically assumes that R&D cost are variable and input prices are taken as given by the firm, which implies $\varepsilon_C = 1$. Furthermore, the R&D production function is assumed to be common across firms and with constant scale elasticity. Taken together, these assumptions imply that the return on R&D is constant across firms and equal to the inverse scale elasticity of the R&D production function.[15]

Note that this insight is unaffected by heterogeneity in the value of R&D V, R&D efficiency expressed as a proportional factor in $F(\ell)$, and input prices as captured by a proportional factor in $C(\ell)$. In fact, firms with larger value of innovation, higher efficiency, and lower input prices conduct more R&D, however, R&D output and cost scale proportionally such that their ratio is constant. This insight is summarized in the Lemma below.

Lemma 1. *Consider a firm solving the maximization problem (3) with production function $F(\ell) = \varphi \cdot \ell^{\frac{1}{1+\phi}}$ and cost function $C(\ell) = W \cdot \ell$. Then, the return on R&D is independent of R&D efficiency φ, the valuation of innovation V, and wages W.*

Proof. Proofs for this section are provided in Appendix C.1. □

In practice, we might expect that at least some of observed variation in R&D return is due to technological differences. A natural source of this variation is industry heterogeneity, which, as discussed above, indeed explains some of the observed dispersion . Nonetheless, it appears unlikely that technology differences are a major driver, since most of the dispersion is among firms in the same industry. Furthermore, my estimates in Appendix B.2 suggest that permanent firm differences are not a major contributor to the observed dispersion.

In absence of technological differences, we can reconcile model with data by allowing for frictions in the choice of R&D investment levels. In this case, the expected return on R&D measures of the degree to which friction prevent firms from choosing their first-best R&D input level. Constrained firms forgo investment projects at the margin with the lowest return within their portfolio and, as a result, have comparatively high average returns, and vice versa. Dispersion in the return on R&D thus becomes a measure of the degree to which frictions affect the allocation of R&D resources across firms.

[15] This applies to the models covered in the most recent literature reviews including Aghion et al. (2014) and Gancia and Zilibotti (2005) as well as more recent contributions including Acemoglu et al. (2018); Akcigit and Kerr (2018); Peters (2020); de Ridder (2021) and Aghion et al. (2022b).

Lemma 2. *Consider a firm subject to friction* Δ *s.t. its first-order conditions satisfy*

$$\frac{\partial F(\ell)}{\partial \ell}\Big|_{\ell=\ell^*} \times V = (1+\Delta) \times \frac{\partial C(\ell)}{\partial \ell}\Big|_{\ell=\ell^*}. \tag{5}$$

Then, its return on R&D is given by

$$\frac{F(\ell^*)V}{C(\ell^*)} = \frac{\varepsilon_C}{\varepsilon_F} \times (1+\Delta). \tag{6}$$

Furthermore, the firms' investment in R&D is decreasing in Δ *such that high returns are associated with low investment in R&D all else equal.*

In companion paper Lehr (2022) I build on this result in an endogenous growth model with fixed inventor supply. Under these assumption, the economic growth rate is given by the product of the frontier growth rate and an allocative efficiency term, which analytically measures the fraction of first-best growth achieved.[16] The latter is determined by frictions Δ and declines in their dispersion. Under the assumption that dispersion in the return on R&D is due to frictions, I estimate an allocative efficiency of around 70% implying a frontier growth rate $1/0.70 - 1 \approx 40\%$ larger than the realized growth rate. The estimate thus suggests that the potential growth impact associated with the measured return on R&D dispersion is large.

Note that expected return on R&D equalization is not necessarily efficient from a growth perspective. Planner and firm incentives differ in endogenous growth models due to the intertemporal externality of knowledge creation (Romer, 1990; Aghion and Howitt, 1992). Building on this insight, the recent literature highlights that resource allocation across firms might be inefficient due to heterogeneous gaps between planner and private valuation of innovation (Acemoglu et al., 2018; de Ridder, 2021; Aghion et al., 2022a). This insight is distinct to my results on dispersion in the R&D return, which is concerned with the optimal allocation of R&D resources from a private perspective. However, the private return on R&D at the firm-level is equalized in all of the before mentioned papers, at odds with my findings.

These concerns might also motivate public R&D policy, which, through subsidies, can affect R&D returns. I investigate this possibility in Appendix B.3 and find little evidence that subsidies explain a significant share of the observed dispersion.

[16] See Appendix D for a full exposition and derivation of the main result.

4.2 Investment and Financial Frictions

A large and growing literature documents dispersion in the return on capital and links it to investment and financial frictions.[17] If firms have limited ability to borrow or face very high cost of external finance, they will forgo marginal investment opportunities and earn high returns as a result. As long as firms differ in their investment opportunities and/or access to external finance, this mechanism gives rise to dispersion in the return on capital.

The same rationale may apply to R&D investment, making financial frictions an intriguing candidate mechanism to explaining R&D return dispersion. According to this theory, financially constrained firms limit their R&D expenditure leading to high returns they forgo funding mediocre projects at the margin.[18] I investigate this link in an OLS framework:

$$\text{Expected Return on R\&D}_{it} = \alpha_{j(i)\times t} + \beta \text{Friction Measure}_{it} + \varepsilon_{it}. \tag{7}$$

If financial frictions are quantitatively important, we would expect $\beta > 0$, i.e. constrained firms have high returns, together with a large R^2. I will use the return on capital as my primary measure for financial frictions together with alternative proxies inspired by the literature including (1) a dummy for whether the firm is listed for less than 20 year, since young firms are considered to be more constrained, (2) a dummy for whether the firm is not paying dividends, since foregoing dividend payments is considered to be a sign of financial hardship, (3) the ratio of cash holdings to assets, since more liquid firms are considered to be less financially constrained, and (4) the Kagan-Zingales Index, which is a summary measure (Kaplan and Zingales, 1997; Whited and Wu, 2006; Midrigan and Xu, 2014).

The estimates in Table 2 suggest that financial frictions do not drive dispersion in R&D returns. Firstly, I find a small and insignificant correlation with the return on capital. Firms that appear to be constrained in their capital investment do not systematically also appear to be constrained in their R&D investment. Secondly, firms that are young, foregoing dividends, or with high Kaplan-Zingales Index have lower returns, which is the opposite of what we would expect if returns on R&D were driven by financial constraints. Finally, liquidity has the expected sign, however, the its explanatory power is low. Financial frictions are thus not quantitatively important for R&D returns, which may be surprising in light of a growing

[17] Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) first documented that there appears to be large dispersion in the return on capital across firms, especially so in developing countries. Asker et al. (2014); Midrigan and Xu (2014); David et al. (2016); David and Venkateswaran (2019); David et al. (2021) link this dispersion to investment and financial frictions.

[18] I formalize the link between R&D returns and financial frictions in Appendix C.1.

literature arguing that R&D investments are especially vulnerable to them (Brown et al., 2009; Peters and Taylor, 2017; Ewens et al., 2020).[19] My evidence thus adds context to the research on the interaction of financial frictions and intangible capital investments.

Table 2: Return on R&D and Measures of Investment Frictions

	(1)	(2)	(3)	(4)	(5)
			Expected Return on R&D		
Return on Capital	0.047				
	(0.067)				
Young Firm		-0.246**			
		(0.098)			
No Dividend Payout			-0.187***		
			(0.047)		
Liquidity				-0.047**	
				(0.022)	
Kaplan-Zingales Index					-0.023
					(0.015)
R2	0.001	0.006	0.006	0.002	0.002
Observations	11,844	11,845	11,845	10,635	11,029

Note: This table reports OLS coefficient estimates. "Mature Firm" and "Dividend Payout" are indicators variable for firm age in Compustat of 20 years or more and positive dividend payments respectively. Liquidity measures the firms cash holdings relative to its book assets. Return and liquity are measured in logs. All regressions control for NAICS3× Year effects and standard errors are clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

4.3 Labor Market Frictions

R&D is mostly a "people business" driven by high-skilled inventors and scientists. According to the 2019 NSF Business Enterprise Research and Development Tables, labor accounts for c. 80% of directly attributable R&D cost.[20] When looking for frictions, it is thus natural to consider the labor market. One friction that recently has received significant attention in the

[19] Relatedly, David et al. (2021) investigate the importance of heterogeneous risk loadings on investment decisions arguing that high risk firms have high required returns. In Appendix B.3 I find that innovation specific risk factor are correlated with the return on R&D, while general risk factors are not.

[20] I calculate total attributable R&D cost as the sum of expenditure on labor, materials, depreciation, and cost of capital. I impute the cost of capital as 1/3 of depreciation, which is in line with a 15% depreciation rate and a 5% cost of capital.

literature is monopsony power.[21] If a firm's labor demand affects its wages, it has an incentive to lower its labor demand to reduce wages, the extend of which depends on the labor supply elasticity.[22] The process of reducing inventor demand, and thereby employment, also forces firms to be more selective about their R&D projects at the margin, which translates into higher R&D returns. Heterogeneity in the firm-specific inventor supply elasticity can thus lead to dispersion in R&D returns as firms facing a lower elasticity reduce their inventor demand more aggressively and, thus, have higher returns.[23]

Dispersion in the return on R&D can thus be explained by heterogeneity in the labor supply elasticity, however, there is no direct evidence thereof for inventors yet. In the following, I will argue that the dispersion in R&D returns is partly driven by monopsony power. In particular, I estimate the inverse labor supply elasticity for inventors and find that it is indeed larger for firms with high returns on R&D.

One approach to estimating the inverse labor supply elasticity is to regress log changes in the inventor wage on changes in log inventor employment at the firm level as shown in equation (8). The coefficient on the changes in inventor employment identifies the average inverse labor supply elasticity if the error term is uncorrelated with changes in inventor employment.

$$\Delta \ln \text{Inventor Wage}_{it} = \bar{\epsilon} \times \Delta \ln \text{Inventors}_{it} + \alpha_{j(i)\times t} + \varepsilon_{it} \tag{8}$$

A natural challenge in this regression are labor supply shocks that simultaneously affect wages and employment. For example, if a firm becomes more attractive to employees for independent reasons, we might expect that the firm will be able to lower wages and higher more workers, however, this variation does not answer the questions as to what happens to wages if the firm wants to expand employment. In other words, supply shocks confounding the estimation of a supply elasticity, and we thus need demand shocks for identification.

Following Seegmiller (2021), I propose to use stock market returns as an instrument for inventor employment. The idea behind the instrument is that stock market returns reflect changes in firm productivity or demand for a firm's product that incentivize it to expand.

[21]See e.g. (Card et al., 2018; Seegmiller, 2021; Berger et al., 2022; Yeh et al., 2022)

[22]There is ample anecdotal evidence that large Tech firms had agreements between each other not to poach employees. Apple, Adobe, Intel, and Google got fined by the Department of Justice in 2010 for illegal non-poaching agreements to keep salaries for tech workers low with further subsequent investigations. See here, here, here. Microsoft only recently announced that it will not enforce its non-compete clauses for employees and was previously sued for their non-poaching agreements. Similar cases have emerged in other industries.

[23]I formalize the link between the labor supply elasticity and R&D returns in Appendix C.1.

Seegmiller (2021) uses the instrument for employment in general, but the argument extends to R&D workers. The identification assumption is thus that stock market returns do not affect changes in inventor wages other than through their impact on the demand for inventor employment.

I connect the inverse labor supply elasticity with the return on R&D by adding an interaction term for firms with above median return on R&D in the above regression framework. If part of the dispersion in the return on R&D is driven by heterogeneous labor supply elasticities, then we would expect a positive coefficient on the interaction term as firms with high returns on R&D face low labor supply elasticities.

$$\begin{aligned}\Delta \ln \text{Inv. Wage}_{it} = \quad & \epsilon_l \times \Delta \ln \text{Inv.}_{it} \\ & + (\epsilon_h - \epsilon_l) \times \Delta \ln \text{Inv.}_{it} \times \{\text{Above Median Return on R\&D}\}_{it} \\ & + \beta\{\text{Above Median Return on R\&D}\}_{it} + \alpha_{j(i)\times t} + \varepsilon_{it}\end{aligned} \tag{9}$$

I measure inventor wages as the ratio of R&D spending to inventors at the 5-year level. In the context of my regression, this is a valid proxy for true inventor wages unless changes in the labor intensity of R&D or share of R&D workers identified by patents are correlated with stock market returns. Using 5-year windows allows me to pick up medium run effects. Note, however, that the instrument only captures annual variation, which safeguards the estimated coefficient from concerns around the use of long-run averages.

My estimation results, as reported in Table 3, reveal two novel findings: first, estimated inverse labor supply elasticities are significantly different from 0 such that expanding firms face higher wages. A 1% increase in employment is associated with a 0.9% increase in average wages. The effect size is of comparable magnitude to Seegmiller (2021)'s estimate for high-skill workers, which is 0.84%. Second, these effects are stronger for firms with high R&D returns. A firm with above median R&D return faces an inverse labor supply elasticity of $0.756 + 1.119 \approx 1.86$ implying that a 1% increase in employment requires a 1.86% increase in wages. Differences in the inverse supply elasticity can account for $\frac{1.119}{2.584-0.848} \approx 64\%$ of the average return difference between below and above median R&D return firms.

I report the first stage results in Appendix B.4. The coefficient have the expected sign and the first stage F-statistic indicates a comfortably high level of power for my instruments. I also report regressions with additional control variables as in Seegmiller (2021). Adding these increases the estimated coefficients significantly and improves their precision.

Table 3: Inventor Inverse Labor Elasticity Estimates

	(1)	(2)
	$\Delta \ln \texttt{Inventor Wage}_{it}$	
$\Delta \ln \texttt{Inventors}_{it}$	0.923***	0.756**
	(0.195)	(0.313)
— $\times$ {Top 50% R&D Return$_{it}$}		1.119**
		(0.496)
{Top 50% R&D Return$_{it}$}		-0.219***
		(0.043)
First stage F stat. (Main)	101	41
First stage F stat. (Inter.)		61
Observations	14,907	14,907

Note: Controls include lagged inventor wage and employment growth as well as current inventor productivity growth. All regressions control for NAICS3 $\times$ year fixed effects. Standard errors clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

5 Monopsony, R&D Return Dispersion, and Growth

I quantify the potential importance of heterogeneous inventor labor supply elasticities for R&D return dispersion and economic growth in a Schumpeterian endogenous growth model with heterogeneous firms. Firms differ in their R&D productivity and labor supply elasticity. I allow for adjustment cost as a stand-in for frictions to labor mobility. I parametrize the model using a combination of external calibration and moment matching.

5.1 Model Description

Time is discrete, infinite and indexed by $t = 0, ..., \infty$. At any point in time there is a constant mass of firms normalized to 1.

Households. The representative household has logarithmic preferences over per-capita consumption c_t and discounts the future with discount factor β. The household consists of a unit mass of workers, whereof a share L are inventors and the remainder production workers. Inventors have labor disutility $u(\{\ell_{it}\})$ depending on their distribution over firms and earn firm-specific wages W_{it}, while production workers have no labor disutility and earn $\tilde{W}_t$. The household owns all firms and earns their profits Π_t. Finally, there is a riskless bond B_t available in zero net-supply paying interest R_t. The household's problem is given by

$$\begin{aligned}
&\max \sum_{t=0}^{\infty} \beta \left(\ln c_t - u(\{\ell_{it}\}) \right) \\
&\text{s.t.} \quad C_t = R_t B_t - B_{t+1} + \int_0^1 \ell_{it} W_{it} di + (1-L)\tilde{W}_t + \Pi_t \\
&\text{and} \quad \int_0^1 \ell_{it} di \leq L.
\end{aligned} \tag{10}$$

I define the labor disutility function implicitly by assuming a labor supply in the spirit of Card et al. (2018) and Kline et al. (2019):

$$\frac{\ell_{it}}{L} = \left(\frac{W_{it}}{W_t} - \bar{\ell} \right)^{\frac{1}{\xi}} \tag{11}$$

The term W_t is a common wage-shifter determined by labor market clearing. The parameter ξ governs the average labor supply elasticity such that we recover the case with common wages and fully elastic supply by setting $\xi = 0$. Intercept parameter $\bar{\ell}$ and "relative wage" $\frac{W_{it}}{W_t}$ determine labor supply across firms. Importantly, this formulation delivers a non-homothetic labor supply elasticity if $\bar{\ell} > 0$, which will be essential to creating dispersion in R&D returns. In Card et al. (2018), ξ is linked to the relative importance of non-monetary compensation and $\bar{\ell}$ to workers' outside option. I discuss the micro-foundation in Appendix C.3.

The standard Euler equation requires

$$\frac{c_{t+1}}{c_t} = \beta R_{t+1}. \tag{12}$$

Static production. Aggregate output Y_t is produced from product-line output y_{jt} with Cobb-Douglas production function. Product-line output is the aggregate across firm production of the particular product, where output across firms are perfect substitutes.

$$\ln Y_t = \int_0^1 \ln(y_{jt}) dj \qquad \text{with} \qquad y_{jt} = \int_0^1 y_{ijt} di. \tag{13}$$

Each firm has a productivity portfolio $\mathcal{A}_{jt} = \{A_{ijt}\}_{j\in[0,1]}$ and produces with linear technology in production labor l_{ijt}:

$$y_{ijt} = A_{ijt} l_{ijt}. \tag{14}$$

Production labor is hired at common production wage $\tilde{W}_t$ and firms compete in Bertrand competition in the product market. Let A_{jt} be the highest productivity level in a product line and $\bar{A}_{jt}$ the second highest, i.e. $A_{jt} = \max_i\{\{A_{ijt}\}_{i\in[0,1]}\}$ and $\bar{A}_{jt} = \max_i\{\{A_{ijt}\}_{i\in[0,1]} \setminus \{A_{jt}\}\}$, then standard limit result is that the firm with productivity $\bar{A}_{jt}$ becomes the sole producer in product line, while charging the marginal cost of second best firm. This equilibrium ensures maximal profits for the best firm without giving any other firm an incentive to produce. Define the leaders productivity advantages as $\lambda_{jt} = A_{jt}/\bar{A}_{jt}$, then equilibrium profits π_{jt}, labor demand l_{jt}, and output y_{jt} are given by

$$\pi_{jt} = Y_t \cdot (1 - 1/\lambda_{jt}), \qquad l_{jt} = \frac{1}{\lambda_{jt}} \frac{Y_t}{\tilde{W}_t} \qquad \text{and} \qquad y_{jt} = A_{jt} l_{jt}. \tag{15}$$

The wage $\tilde{W}_t$ is pinned down by market clearning:

$$1 - L = \int_0^1 \ell_{jt} dj. \tag{16}$$

As shown in Peters (2020), this setup gives rise to a simple aggregate production function depending on aggregate productivity index A_t, a production efficiency term Λ_t depending on the distribution of $\{\lambda_{jt}\}$, and the mass of production workers L:

$$Y_t = A_t \Lambda_t (1 - L) \quad \text{with} \quad \ln A_t = \int_0^1 \ln(A_{jt}) dj \quad \text{and} \quad \Lambda_t = \frac{\exp\left(\int_0^1 \ln\left(\frac{1}{\lambda_{jt}}\right) dj\right)}{\int_0^1 \left(\frac{1}{\lambda_{jt}}\right) dj}. \tag{17}$$

Innovation. Firms innovate to become leaders in new product lines. In turns, they lose their status as a leader whenever a competitor innovates in one of their product lines. Firms innovate with probability z_{it} depending on their R&D efficiency φ_{it} and hired inventors ℓ_{it}:

$$z_{it} = e^{\mu + \varphi_{it}} \ell_{it}^{\frac{1}{1+\phi}}, \tag{18}$$

where μ governs the common R&D productivity. When a firm successfully innovates, it becomes the leader in a random new product line and draws associated productivity advantage $\lambda \sim f_\lambda$. I assume that the distribution over potential productivity advantages is common across firms and constant over time.

The firm's idiosyncratic R&D efficiency follows an AR(1) process:

$$\varphi_{it} = \rho\varphi_{it-1} + \nu_{it} \qquad \nu_{it} \overset{i.i.d.}{\sim} N(0, \sigma_\varphi^2). \tag{19}$$

Firms face R&D cost $C(\ell_{it}, \ell_{it-1})$, which features heterogeneous, finite labor supply elasticities via a firm specific innovator wage W_{it} as well as adjustment cost AC_{it}:

$$C_t(\ell_{it}, \ell_{it-1}) = W_{it}\ell_{it}(1 + AC_{it}). \tag{20}$$

The labor supply formulation gives rise to the firm specific innovator wages with finite and heterogeneous labor supply elasticity, which firms take into account:

$$W_{it} = W_t\left((\ell_{it})^\xi + \bar{\ell}\right). \tag{21}$$

I allow for quadratic adjustment cost to allow for dynamic frictions in the labor market:

$$AC_{it} = \gamma\left(\frac{\ell_{it} - (1-\delta)\ell_{it-1}}{\ell_{it-1}}\right)^2. \tag{22}$$

Here, γ captures the strength of adjustment cost and δ natural employment turnover.

With slight abuse of notation, I will denote a firms productivity advantage portfolio by $\mathcal{A}_{it} = \{\lambda_{jt}\}_{j\in\mathcal{J}_{it}}$, where $\mathcal{J}_{it}$ is the set of product lines in which firm i is the leader at time t. Note that this set could be empty.

The firm's dynamic problem is given by

$$V_t(\mathcal{A}_{it}, \varphi_{it}, \ell_{it-1}) = \max_{\ell_{it}}\left\{\sum_{j\in\mathcal{J}_{it}}\pi_{jt} - C_t(\ell_{it-1}, \ell_{it}) + \left(\frac{1}{R_t}\right)\mathbb{E}_t\left[V_{t+1}(\mathcal{A}_{it+1}, \varphi_{it+1}, \ell_{it})\right]\right\}. \tag{23}$$

Expectations are taken with respect to the R&D efficiency process, a potential realization of λ, and the evolution of the existing product lines in $\mathcal{A}_{it}$. For each $j \in \mathcal{J}_{it}$, the firm remains the leader with probability $1 - z_t$ and loses its leader status otherwise, where z_t is the aggregate innovation rate:

$$z_t = \int_0^1 z_{it}di. \tag{24}$$

Thus, for $\lambda_{jt} \in \mathcal{A}_{it}$, $\lambda_{jt} \in \mathcal{A}_{it+1}$ with probability $1 - z_t$. Furthermore, with probability z_{it} a λ drawn from distribution f_λ becomes part of $\mathcal{A}_{it+1}$.

The common wage-shifter W_t is determined by labor market clearing for inventors:

$$L = \int_0^1 \ell_{it} di. \tag{25}$$

Definition 1. *A competitive equilibrium is a sequence of prices* $\{W_t, R_t\}$, *quantities* $\{\ell_{it}, Y_t, A_t, \Lambda_t\}$, *productivity portfolios* $\{\mathcal{A}_{it}\}$ *and efficiency distributions* $\{\varphi_{it}\}$, *and value function* $\{V_t(\mathcal{A}_{it}, \varphi_{it}, \ell_{it-1})\}$ *such that firms optimize, markets clear, and the above defined laws of motion are satisfied.*

5.2 Characterizing the Equilibrium

I analyze, calibrate, and simulate the model in recursive formulation along a balanced growth path, which is summarized in Definition 2. Two properties are useful in deriving the recursive form. First, the model formulation allows me to decompose the value function into a profit and R&D component. The former captures the expected net-present-value of the profits associated with existing leadership positions and is independent of a firm's R&D choices. The latter captures the value of the firm's ability to conduct R&D and create leadership positions in the future. Second, as in other endogenous growth models, the growth-rate in this economy is the expected productivity improvement of an invention times the aggregate innovation rate. The latter crucially depends on the allocation of inventors and is the direct vehicle through which frictions impact growth. See Appendix C.2 for details.

Definition 2. *A recursive Balanced Growth Path equilibrium is a growth rate* g, *prices* $\{W, R, \mathcal{V}(\lambda)\}$, *value function* $\tilde{V}(\varphi, \ell)$ *with policy function* $\ell'(\varphi, \ell)$, *distribution* f_λ, f_φ *and* $f(\varphi, \ell)$ *such that*

- *the value function and policy function solve*

$$\begin{aligned} V(\ell, \varphi) &= \max_{\ell'} \{-C(\ell, \ell') + \beta\left(z(\varphi, \ell')\mathbb{E}_\lambda[\mathcal{V}(\lambda)] + \mathbb{E}[V(\ell', \varphi')]\right)\} \\ s.t. \quad C(\ell, \ell') &= W\ell'\left(\ell' + \bar{\ell}\right)^\xi \left(1 + \gamma\left(\frac{\ell'}{\ell} - (1-\delta)\right)^2\right) \\ \mathbb{E}_\lambda[\mathcal{V}] &= \frac{1 - \mathbb{E}_\lambda[1/\lambda]}{1 - \beta(1-z)} \quad \text{and} \quad z(\varphi, \ell') = e^{\mu+\varphi} \cdot \ell'^{\frac{1}{1+\phi}}. \end{aligned} \tag{26}$$

- *the aggregate innovation and growth rate are given by*

$$g = z \cdot \mathbb{E}_\lambda[\ln \lambda] \quad \text{and} \quad z = \int \left(e^{\mu+\varphi} \cdot \ell'(\varphi, \ell)^{\frac{1}{1+\phi}}\right) dF(\varphi, \ell) \tag{27}$$

- *labor market clearing holds*

$$\mathcal{L} = \int \ell'(\varphi, \ell) dF(\varphi, \ell) \tag{28}$$

- *the distribution function $f(\varphi, \ell)$ satisfies*

$$f(\varphi', \ell') = \int f(\varphi, \ell)\{\ell'(\varphi, \ell) = \ell'\} f_\varphi(\varphi'|\varphi) dF(\varphi, \ell), \tag{29}$$

where $f_\varphi(\varphi'|\varphi)$ is the conditional density over φ'.

The model generates dispersion in the return on R&D through two channels: Non-homothetic wages ($\bar{\ell} > 0$) and adjustment cost ($\gamma > 0$). Adjustment cost lead to dispersion as firms adjust their R&D expenditure gradually in response to a R&D productivity shocks, while R&D output responds immediately. Thus, firms receiving a positive productivity shock have temporarily elevated R&D returns and vice versa.

Lemma 3. *Suppose $\gamma = 0$, then the expected return on R&D is given by*

$$\frac{z(\varphi, \ell^*(\varphi))\mathbb{E}_\lambda[\mathcal{V}(\lambda)]}{C(\ell^*(\varphi))} = (1 + \phi) \cdot \left(1 + \xi \cdot \frac{\ell^*(\varphi)}{\ell^*(\varphi) + \bar{\ell}}\right). \tag{30}$$

Non-homothetic wages, as summarized in Lemma 3, yield dispersion in R&D returns due to local differences in the labor supply elasticity. Firms with large inventor employment face less elastic supply and thus have stronger incentives to suppress their inventor demand in order to reduce their wages. The formulation thus connects the labor supply elasticity with labor demand and, in equilibrium, with R&D productivity. Three factors push me towards such a specification. Firstly, it is in line with the evidence in Seegmiller (2021), who shows that labor supply elasticities are lower with very productive firms, and especially so among highly productive workers.[24] Secondly, I show in Appendix B.3 that more dominant firms in the inventor market indeed have higher returns. Finally, as discussed in Appendix C.5, this formulation is in line with the finding that the aggregate is larger than the average return on R&D, which requires a positive correlation between R&D returns and expenditure.

[24] Note that the monopsony model in Berger et al. (2022) also has the feature that larger firms face more inelastic labor supply.

5.3 Numerical Solution and Simulation

I solve the model numerically using discretization methods. I create a large grid for labor input choices and discretize the productivity process using the Tauchen method. I then solve the firm's problem via value function iteration and employ non-stochastic simulation to calculate aggregates for market clearing. My baseline algorithm enforces a growth rate of 1.5% p.a. via the average R&D efficiency parameter μ. See Appendix C.4 for further details.

I calculate model moments by simulating a single firms for 100,000 periods with an additional 50 "burn-in" periods at the beginning. I assume that the firm has N_P different R&D lines with perfectly correlated R&D productivity process. R&D success and patent valuation are independent across product lines making the number of patents per period a Bernoulli random variable. I add ex-ante uncertainty in patent valuations and, thus, ex-post measurement error in R&D returns by drawing them from a geometric distribution:[25]

$$\lambda_{it} = \lambda^{\Delta_{it}} \quad \text{with} \quad Pr(\Delta_{it}) = (1-P)P^{\Delta-1} \quad \text{for} \quad \Delta = 1, 2, ... \tag{31}$$

Using the simulated data I construct the relevant sample moments. Throughout, I perform the same operations on the simulated data as for deriving my empirical estimates.

5.4 Calibration

I parameterize the model with a combination of external calibration and moment matching.

External calibration. Firstly, I set the discount factor β to 0.97, which, together with a targeted growth rate of 1.5%, implies an annual interest rate of c. 4.5% and is broadly in line with standard calibrations (Acemoglu et al., 2018). Secondly, I set the R&D scale elasticity ϕ to 1 (Acemoglu et al., 2018). The elasticity controls firms' sensitivity to R&D productivity shocks. Finally, I calibrate the depreciation rate for R&D workers δ to 12.5%, which matches the natural turnover of employees in the LED Quarterly Workforce Indicators.[26] Higher levels of δ lead to an asymmetry in adjustment cost, making it costlier to grow than to shrink.

[25] Note, however, that the firm is risk-neutral and, thus, only takes the expected value of profits $\mathbb{E}[\pi(\lambda_{it})]$ into consideration.

[26] In the data, I first calculate the turnover of employees that is not linked to net-flows as gross minus net worker turnover, which captures workers turnover within industries, and then normalize this number by employment and take a simple average to get an aggregate estimate of 12.5%.

Internal calibration/ moment matching. I split the internal calibration into two steps. Firstly, I calibrate the parameters for the patent valuation process, λ and P, to match an average markup of 20% together with the within firm-year standard deviation of log patent valuations in my sample. The step-size parameter λ primarily controls the average markup, while P is closely linked to the dispersion of patent valuation. I can perform this step separately as both moments only depend on the process for λ_{it}. Imposing a relatively large λ via a large average markup ensures that the probability of invention z remains well below 1. Note that conditional on a targeted growth rate, the size of λ does not influence the aggregate as larger values simply imply lower required levels of average R&D efficiency.

After this step, five parameters remain to calibrate: the standard deviation σ and auto-correlation ρ of the R&D efficiency process, the parameters of the wage function, ξ and $\bar{\ell}$, and adjustment cost γ. I calibrate them by targeting the 8 moments listed in Table 4, which concern the behavior of R&D expenditure and inventor employment, estimated wage elasticities, and auto-correlations of R&D returns. I match moments by minimizing the weighted distance of model and data moments using absolute differences in percent except for the auto-correlation for R&D growth, where I use level differences. Moments are weighted to emphasize the basic behavior of R&D expenditure together with my estimates from the previous section. The targeted moments and parameters are intimately linked in the model, however, some relationships are particularly important for identification. Firstly, the standard deviation of R&D growth is positively linked to the dispersion in R&D productivity shocks. Secondly, the auto-correlation of R&D productivity and adjustment costs both increase the auto-correlation of log R&D and its changes, however, with different sensitivities. Finally, the wage elasticity estimates are linked to the wage function, where the average elasticity is primarily governed by ξ, while the relative elasticities allow us to identify $\bar{\ell}$ by governing the dispersion in R&D returns. More heterogeneity generally requires larger $\bar{\ell}$.

Table 4 reports the targeted moments together with their model counterparts. The model fits well overall, but does not quite match the heterogeneity in the supply elasticity as captured by moments 7 and 8. The model also delivers auto-correlated R&D returns, however, it does not quite capture the magnitude. Auto-correlation arises due to combination of the wage function and auto-correlated R&D productivity. In particular, the wage function links the inverse labor supply elasticity, and thus R&D returns, to the labor demand. The latter is positively auto-correlated due to the productivity process, making the return on R&D auto-correlated as well.

Table 4: Model vs Data Moments

Moment	Data	Model	Target	Source
Average markup	0.200	0.200	λ	Norm.
SD of log patent valuations	0.562	0.562	P	Data
SD of R&D growth	0.316	0.316	σ	Data
Auto-corr. of log R&D	0.922	0.922	ρ	Data
Auto-corr. of R&D growth	-0.017	0.028	γ	Data
Avg. wage elasticity	0.923	0.986	$\{\xi, \bar{\ell}\}$	Data
Avg. wage elas. for low R&D returns	0.756	0.672	ξ	Data
Diff. in avg. wage elas. high vs low R&D returns	1.119	1.089	$\bar{\ell}$	Data
Inventor - R&D expenditure elas.	0.638	0.517	$\{\xi, \bar{\ell}\}$	Data
Auto-corr. of Return on R&D	0.651	0.437	$\{\xi, \bar{\ell}, \gamma\}$	Data

Note: This table reports model and data moments targeted in the calibration the model with monopsony power. Model values based on simulation with 100,000 observations. I estimate the auto-correlations accounting for permanent firm differences as in Han and Phillips (2010). The estimated wage elasticities respond to the estimates in columns (1) and (2) of Table B.13. The return on inventors is defined as the ratio of patent valuations to inventors. The final two auto-correlations are calculated at the 5-year horizon.

Table 5 reports the calibrated parameters. R&D productivity is highly auto-correlated and its innovations are highly volatile.[27] The large calibrated volatility is mainly due to the presence of monopsony power, which reduces R&D expenditure volatility by increasing the concavity of the firm's objective function. The calibrated adjustment cost are small and imply that a firm increasing its employment by 10% faces an additional cost of 0.4% of its wage bill.[28] Finally, the calibration for labor supply is best understood through its effect on the wage and their elasticity. Under the calibration, both are highly convex. Firms with low employment face a wage elasticity close to 0, which effectively makes them price-takers. On the other hand, firms with large R&D employment face an inverse labor supply elasticity in excess of 2, such that raising employment by 1% requires an increase of the average wage by 2%. This difference has large effects on wages. Firms at the upper end of the employment distribution pay workers around 3 times as much as low R&D employment firms.

[27] For example, the volatility of profitability shocks, which directly map into R&D productivity, in Terry (2022) is about one fourth of my parameter estimate, while the auto-correlation is of comparable magnitude.

[28] Cooper and Haltiwanger (2006) estimate an adjustment cost parameter of 0.455 for capital investment, while Asker et al. (2014) estimate a value above 8.

Table 5: Calibrated Parameters

Parameter	Description	Value	Source
β	Discount factor	0.970	External calibration
ϕ	R&D scale elasticity	1.000	Acemoglu et al. (2018)
L	Researchers	0.142	Acemoglu et al. (2018)
δ	Inventor turnover	0.120	External calibration
λ	Minimum step size	1.080	Internal calibration
$\bar{P}$	Step size shape parameter	0.447	Internal calibration
σ	Std. dev. R&D prod. shocks	0.446	Internal calibration
ρ	Auto-corr. R&D prod.	0.867	Internal calibration
γ	Adjustment cost	0.101	Internal calibration
ξ	Avg. monopsony	4.755	Internal calibration
$\bar{\ell}$	Rel. monopsony	3.884	Internal calibration

Note: Table reports model calibration.

Data- vs model-implied returns. The model links R&D returns to inventor employment via the wage elasticity. I test this link empirically by estimating the log-transformed expression in Lemma 3 via OLS, constructing the model-implied wage elasticity by combining the calibrated parameters of the wage function with the empirical distribution of inventors. The coefficient estimate in column (1) of Table 6 confirms a strong positive relationship between the return on R&D and the estimated wage elasticity, explaining about 9% of the variation with a coefficient estimate around 0.4.[29] Furthermore, column (3) confirms that this link is not driven a general correlation between inventor employment and R&D returns. This exercise, thus, suggests that the calibrated wage formulation indeed captures the relationship between inventor employment and R&D returns, and that the associated labor market imperfections are indeed a driving forces behind the dispersion in R&D returns.

[29] Note that the coefficient estimate is about one third the model implied elasticity. The difference could be driven by other factors influencing R&D returns or measurement error in inventor employment.

Table 6: Return on R&D and Implied Wage Elasticity

	(1)	(2)	(3)
		ln **Return on R&D**	
$\ln(1+\hat{\varepsilon}_W)$	0.383***		0.304***
	(0.038)		(0.081)
ln Inventors		0.234***	0.056
		(0.032)	(0.065)
R2	0.08	0.07	0.08
Observations	11,812	11,812	11,812

Note: This table reports OLS regression coefficients. The implied inverse supply elasticity is estimated using inventor employment and calibrated parameters. Standard errors are clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

5.5 Results

The calibrated model provides an estimate of the contribution of heterogeneous monopsony power to R&D return dispersion. As reported in the first row of Table 7, the calibrated model delivers a standard deviation of log returns on R&D of around 0.35, which is around 1/3 of its standard deviation in the data.[30] Forcing firms to take wages as given reduces the dispersion to close to 0 as shown in the second row, while turning-off adjustment costs has a negligible impact as show in the third row. Jointly, this suggests that heterogeneous monopsony power is a meaningful contributor to the measured dispersion, but, nonetheless, leaves much variation to be explained.

The model suggests that the documented frictions have a significant impact on economic growth. As reported in the second row of Table 7, annual growth would be 0.06 p.p. (4%) larger in the model if firms did not take into account their pricing power. Imposing no dispersion in the return on R&D by also turning-off adjustment cost increases this cost to 0.08 p.p. (5%). The associated welfare cost amount to 2.1% and 2.8% in consumption-equivalent terms, respectively. For comparison, Lucas (2003) estimates that the cost of business cycles are around 1%, while Arkolakis et al. (2012) argue that US welfare would decrease by

[30] I only report values for dispersion in expected returns. Realized returns have a small measurement error component in the model, which is unrelated to the frictions driving dispersion in expected returns.

1% under trade autarky. In the misallocation literature, Hsieh and Klenow (2009) estimate welfare cost of capital misallocation in the US around 30-40%, while Berger et al. (2022) estimate welfare cost of monopsony power in production at 7.6% due to 20.9% lower output.

Table 7: Return Dispersion, Growth, and Monopsony

Model	SD	Growth-rate	Welfare
Baseline	0.35	1.50%	—
No monopsony	0.03	1.56%	2.1%
No adjustment cost	0.34	1.50%	0.1%
No dispersion	0.00	1.58%	2.8%

Note: Table reports model results for main calibration and counterfactual where firms take wages as given. SD refers to the standard deviation of log R&D returns based on simulation with 100,000 periods. Welfare column quantifies growth-rate change in terms of consumption equivalent change.

It should be emphasized that the documented cost are driven by the impact of friction on the allocation of inventors across firms only as the number of inventors is fixed. These frictions thus reduce growth by decreasing the allocative efficiency in the R&D sector.

Discussion - Perfect price discrimination. Monopsony power arises when the firm's marginal hiring decision has an impact on its inframarginal wage. In contrast, under perfect price discrimination, firms' marginal hiring decision do not affect the wages of other workers and, thus, firms have no incentive to artificially keep their labor demand low. I show in Appendix C.6 that R&D return dispersion can still arise under price discrimination, but its source is different. Price discrimination breaks the proportionality between the average and the marginal wage leading to dispersion in the average Return on R&D, but not the marginal one. Thus, R&D return dispersion itself might not be a sign of misallocation if we have strong reason to believe that marginal price and benefits faced by the firms are not proportional to their averages, however, the evidence in Seegmiller (2021) suggests that marginal hiring decision for high-skilled workers do affect the wage of the existing workforce.[31]

[31] Note that a similar insight applies to the welfare cost of measured markups and is at the heart of some of the findings in Bornstein and Peter (2022).

6 Conclusion

I document that R&D returns are widely dispersed among US listed firms. This is surprising as workhorse models of endogenous growth predict that R&D resources should flow from low to high return firms until returns are equalized. Dispersion then suggests a failure to allocate its R&D resources efficiently and, as a result, less innovation and growth.

I investigate frictions as a potential mechanism and find mixed results. While financial frictions appear unrelated to the documented dispersion, monopsony power in the market for inventors appears to be an important force. I estimate that firms with high return on R&D have stronger pricing power in the market for inventors such that they can keep wages low by hiring fewer of them and, as a by-product, achieve higher average returns.

I quantify the contribution of monopsony power in a quantitative Schumpeterian growth model, which I calibrate to match key estimates and data moments. The calibrated model accounts for 1/3 of the standard deviation of the Return on R&D and predicts a 0.06 p.p. larger growth rate if firms did not take into account their labor market power, equivalent to a consumption-equivalent welfare gain of 2.1%. Dispersion in the return on R&D is thus has important implications for economic growth and welfare.

Jointly, my findings suggest as least three avenues for future research. Firstly, a large share of the R&D return dispersion remains unaccounted for and, thus, future research is needed to explore its underlying sources. Secondly, this paper highlights the growth impact of monopsony power over inventors. Further research on other inventor market frictions and their impact the innovation ecosystem is required. Finally, more research is needed to understand in how far policy can tap into the large growth potential implied by the documented dispersion in R&D returns.

References

Acemoglu, Daron and Dan Cao, "Innovation by entrants and incumbents," *Journal of Economic Theory*, 2015, *157*, 255–294.

__ , Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William R. Kerr, "Innovation, Reallocation, and Growth," *American Economic Review*, 2018, *108*, 3450–3491.

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60*, 323–351.

__ , Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li, "Good Rents versus Bad Rents: R&D Misallocation and Growth," 2022.

__ , __ , __ , Peter J. Klenow, and Huiyu Li, "A Theory of Falling Growth and Rising Rents," 2022.

__ , Ufuk Akcigit, and Peter Howitt, *What Do We Learn From Schumpeterian Growth Theory?*, Vol. 2, Elsevier B.V., 2014.

Akcigit, Ufuk and William R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126*, 1374–1443.

__ , Douglas Hanley, and Stefanie Stantcheva, "Optimal Taxation and R&D Policies," *Econometrica*, 2022, *90*, 645–684.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New trade models, same old gains?," *American Economic Review*, 2 2012, *102*, 94–130.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122*, 1013–1063.

Azar, José, Steven T. Berry, and Ioana Elena Marinescu, "Estimating Labor Market Power," *SSRN Electronic Journal*, 2019.

Baber, William R, Patricia M Fairfield, and James A Haggard, "The Effect of Concern about Reported Income on Discretionary Spending Decisions: The Case of Research and Development," 1991.

Berger, David W., Kyle Herkenhoff, and Simon Mongey, "Labor Market Power," *American Economic Review*, 2022, *112*, 1147–1193.

Berkes, Enrico, "Comprehensive Universe of U.S. Patents (CUSP): Data and Facts," *Working paper*, 2016, pp. 1–27.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," *American Economic Review*, 2020, *110*, 1104–1144.

_ , Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81*, 1347–1393.

Bornstein, Gideon and Alessandra Peter, "Nonlinear Pricing and Misallocation," 2022.

Brown, James R., Steven M. Fazzari, and Bruce C. Petersen, "Financing innovation and growth: Cash flow, external equity, and the 1990s r&d boom," *Journal of Finance*, 2 2009, *64*, 151–185.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36*, S13–S70.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D investment with corporate income tax cuts in China," *American Economic Review*, 2021, *111*, 2065–2100.

Cohen, Wesley M. and Steven Klepper, "A Reprise of Size and R&D," *The Economic Journal*, 1996, *106*, 458–470.

Cooper, Russell W. and John C. Haltiwanger, "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 2006, *73*, 611–633.

Cunningham, Colleen, Florian Ederer, and Ma Song, "Killer Acquisitions," 2020.

David, Joel M. and Venky Venkateswaran, "The sources of capital misallocation," *American Economic Review*, 2019, *109*, 2531–2567.

_ , Hugo A. Hopenhayn, and Venky Venkateswaran, "Information, Misallocation, and Aggregate Productivity," *The Quarterly Journal of Economics*, 5 2016, *131*, 943–1005.

_ , Lukas Schmid, and David Zeke, "Risk-Adjusted Capital Allocation , Misallocation and Aggregate TFP," 2021.

de Ridder, Maarten, "Market Power and Innovation in the Intangible Economy," 2021.

Dukes, Roland E, Thomas R Dyckman, and John A Elliott, "Accounting for Research and Development Costs: The Impact on Research and Development Expenditures," 1980.

Ewens, Michael, Ryan Peters, and Sean Wang, "Measuring Intangible Capital with Market Prices," 2020.

Gancia, Gino and Fabrizio Zilibotti, "Horizontal Innovation in the Theory of Growth and Development," *Handbook of Economic Growth*, 2005, *1*, 111–170.

Han, Chirok and Peter C.B. Phillips, "GMM estimation for dynamic panels with fixed effects and strong instruments at unity," *Econometric Theory*, 2010, *26*, 119–151.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124*, 1403–1448.

_ **and** _ , "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 2014, *129*, 1035–1084.

Kaplan, Steven N and Luigi Zingales, "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?," *The Quarterly Journal of Economics*, 1997, *112*, 169–215.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation over the Long Run," *American Economic Review: Insights*, 9 2021, *3*, 303–20.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who Profits from Patents? Rent-Sharing at Innovative Firms," *The Quarterly Journal of Economics*, 2019, *134*, 1343–1404.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," *The Quarterly Journal of Economics*, 5 2017, *132*, 665–712.

Lehr, Nils Haakon, "Misallocation and the Growth Slowdown," 2022.

Lerner, Josh, "Patenting in the Shadow of Competitors," *The Journal of Law & Economics*, 1995, *38*, 463–495.

Lev, Baruch, Bharat Sarath, and Theodore Sougiannis, "R&D Reporting Biases and Their Consequences," *Contemporary Accounting Research*, 2005, *22*, 977–1026.

Lucas, Robert E., "Macroeconomic priorities," *American Economic Review*, 3 2003, *93*, 1–14.

Lucking, Brian, "Do R&D Tax Credits Create Jobs?," 2019.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and misallocation: Evidence from plant-level data," *American Economic Review*, 2014, *104*, 422–458.

Peters, Michael, "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 2020, *88*, 2037–2073.

Peters, Ryan H. and Lucian A. Taylor, "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 2017, *123*, 251–272.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 10 2008, *11*, 707–720.

Romer, Paul M, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98.*

Seegmiller, Bryan, "Valuing Labor Market Power : The Role of Productivity Advantages," 2021.

Shi, Liyan, "The Macro Impact of Noncompete Contracts," 2020, pp. 1–41.

Syverson, Chad, "Challenges to Mismeasurement Explanations for the US Productivity Slowdown," *Journal of Economic Perspectives*, 5 2017, *31*, 165–186.

Terry, Stephen J., "The Macro Impact of Short-Termism," 2022.

_ , Toni M. Whited, and Anastasia A. Zakolyukina, "Information Versus Investment," *Review of Financial Studies*, 2022, *(forthcoming).*

Whited, Toni M and Guojun Wu, "Financial Constraints Risk," *The Review of Financial Studies*, 2006, *19*, 531–559.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein, "Monopsony in the U.S. Labor Market," *American Economic Review*, 2022, *112*, 2099–2138.

Appendix

A Data Appendix

A.1 Data Construction

Mapping patents to firms. I assign patents to firms based on the crosswalk between patents and `PERMNO`s in Kogan et al. (2017), which I extend to `GVKEY`s using the mapping provided by WRDS.

Measuring inventor employment. Let $\mathcal{P}_{it\to t+4}$ be the set of successful patent applications for firm i between t and $t+4$ and $\mathcal{I}_{it\to t+4}$ be the set of associated inventors. I will denote the number of patents assigned to firm i and listing j as inventor at time t as P_{ijt} and the total number of patents listing j as inventor as P_{jt}

$$\texttt{Inventors}_{it\to t+4} = \sum_{j\in\mathcal{I}_{it\to t+4}} \frac{\sum_{s=0}^{4} P_{ijt+s}}{\sum_{s=0}^{4} P_{jt+s}}. \tag{A.1}$$

I use two additional measure in robustness checks. Firstly, I use the raw size of $|\mathcal{I}_{it\to t+4}|$, which forgoes the full-time equivalent adjustment, and, secondly, I construct the measure first at the 1-year horizon and then aggregate over the 5-year window. Note that the former is identical to the main measure when all inventors are only listed on patents that are also assigned to the firm.

Inventor wages. I construct inventor wages as the ratio of R&D expenditure and my inventor employment measure.

$$\texttt{Inventor Wage}_{it} \equiv \frac{\sum_{s=0}^{4} \texttt{R\&D Expenditure}_{it+s}}{\texttt{Inventors}_{it\to t+4}}$$

Dominance. I construct the dominance measure used in Appendix B.3 in two steps. Firstly, for each of the firm's new patent within a 5-year window, I calculate the share of inventors working for the firm among those that worked on patents of the exactly same technology class classification. For the latter, I use the complete CPC classification of the patent, which has more than 600 technology classes, which are non-exclusive at the patent level. Patents of the same technology class are thus those that have exactly the same classifications as the patent in consideration. As before, I distinguish between inventors using the USPTO disambiguation and link inventors to a firm if they are listed on a firm's new patent for the 5-year window in consideration.

Secondly, I aggregate the patent-based measure to the firm-level by taking a simple average over the firm's new patents. Note that the resulting measure is between 0 and 1 by construction with 1 implying maximal dominance and vice versa.

Specialization. I construct the specialization measure used in Appendix B.3 in two steps. Firstly, I calculate inventor specialization for a given 5-year window as the average cosine similarity between patent classifications in an inventors portfolio of new patents. I rely on CPC classifications of patents, which has more than 600 non-exclusive patent categories. For each patent I then create an indicator vector over the set of available patent classification, where I weight individual categories by their inverse frequency. I then calculate the average cosine similarity across all patents in the portfolio and take the simple average across all patents. This measure is between 0 and 1 by construction with 0 implying completely different patents and 1 implying that all patents have the same technology classification.

I aggregate this measure up to the firm-level by taking a patent-weighted average across inventor associated with a firm, where the weight reflect the number of new patents shared by the inventor and firm. I interpret a larger value in this measure as more specialized inventors and vice verse following the logic that specialized inventors work on similar patents.

A.2 Constructing Adjusted Returns

I construct adjusted returns on R&D in three steps. Firstly, I construct a baseline return using winsorized patent valuations and knowledge capital for R&D output and input respectively. Winsorizing patent valuations reduces the impact of outliers and thus dispersion in the measured returns. I winsorize the top 5% of valuations in each application year. Knowledge capital is the discounted aggregate of R&D expenditure and, thus, is a more long-term input measure. I use the knowledge capital measure constructed in Ewens et al. (2020).

Secondly, I adjust for the potential impact of acquiring innovative firms with patents in the pipeline. This could lead to measurement error as R&D expenditure is occurred by the firm before acquisition, while patents are realized post acquisition. Let C_{it} be the R&D expenditure, V_{it} the value created from the firm's R&D, A_{it} be the value of acquisitions and

s the share of acquisition that effectively is R&D expenditure. Then measured returns are given by

$$\frac{V_{it}}{C_{it}} = \frac{\tilde{V}_{it}}{C_{it} + sA_{it}} \times \left(1 + s\frac{A_{it}}{C_{it}}\right) \tag{A.2}$$

If firms optimize appropriately, then the first component is a constant such that

$$\frac{\partial \ln(V_{it}/C_{it})}{\partial(A_{it}/C_{it})} \approx s, \tag{A.3}$$

where I assumed that the share of innovation expenditure in acquisition is small. We can then estimate s via OLS and add the R&D part of acquisitions to the knowledge capital input measure.

Finally, I adjust for industry heterogeneous and amenities by residualizing the log return with respect to NAICS3×year effects as well as average temperatures, house prices, and income levels in the location of inventors working for a firm. The latter accounts for amenities that factor into a firm's overhead cost for R&D, but are not reported as R&D expenditure.

B Empirical Appendix

B.1 Robustness for Return on R&D Dispersion

Industry differences. Table B.1 confirms that return dispersion remains large when we only consider variation within industries at particular point in time. For example, the return on R&D dispersion within NAICS3 $\times$ Year cells is 0.93, which is about 85% of the total variation. This finding suggests that R&D dispersion is driven by firm specific, idiosyncratic factors instead of common industry characteristics.

Furthermore, the gap in return dispersion remains large as well when we focus on variation within narrow comparison groups. For example, the Return on R&D has a 45% and 82% larger standard deviation compared to the Return on Capital and Labor respectively when focusing on within NAICS3 $\times$ Year variation.

Table B.1: Return Dispersion Across Comparison Groups

Within Cell	Return on R&D	Return on Capital		Return on Labor	
	SD	SD	$\Delta\%$	SD	$\Delta\%$
—	1.11	0.77	43.4%	0.79	39.6%
Year	1.06	0.74	44.2%	0.60	76.0%
NAICS3 $\times$ Year	0.93	0.64	46.4%	0.51	83.5%
NAICS6 $\times$ Year	0.84	0.58	45.7%	0.45	88.6%

Note: Return measures residualized with respect to fixed effects indicated in first column. Column headers SD report standard deviations of return measure. Columns headers $\Delta\%$ indicate percent difference of Return on R&D dispersion with respect to return in consideration. Returns are measured in logs.

Stochastic realizations. The realization of R&D expenditure might have a stochastic component, which could yield dispersion in measured R&D returns. I investigate this concern by assuming that a share $P_h(p)$ of R&D expenditure is realized at horizon h, where P_h is the geometric distribution:

$$P_h(p) = \frac{(1-p)^{h-1} \cdot p}{1-(1-p)^{\bar{\Delta}}} \quad \text{for} \quad h = 1, ..., \bar{\Delta} \quad \text{with} \quad \bar{\Delta} = 10.$$

The relevant R&D expenditure for innovation at time t is then given by

$$\text{R\&D}^p_{it-1} = \sum_{h=1}^{\bar{\Delta}} \text{R\&D}_{t-s} \cdot P_h(p),$$

and the value associated with R&D investment at time $t-1$ is given by

$$\text{Valuation}^p_{it} = \sum_{h=1}^{\bar{\Delta}} \frac{\text{R\&D}_{it-1} \cdot P_h}{\text{R\&D}^p_{it-1+h}} \cdot \text{Valuation}_{it+h-1}.$$

I then construct alternative measures of R&D return at the 5-year window as

$$\text{Expected Return on R\&D}^p_{it} \equiv \frac{\sum_{s=0}^{4} \text{Valuation}^p_{it+s}}{\sum_{s=0}^{4} \text{R\&D Expenditure}_{it-h+s}}. \tag{B.1}$$

Note that $p = 100\%$ recovers the baseline case.

Table B.2 confirms that the large dispersion of R&D returns and its gap with respect to other measures of return dispersion is highly robust to this alternative specification. For example, a steep, but not instantaneous decay of $p = 95\%$ reduces the dispersion in the Return on R&D marginally, while leaving it more than 40% larger than the Return on Capital. Particularly slow decays increase the dispersion in measured R&D returns.

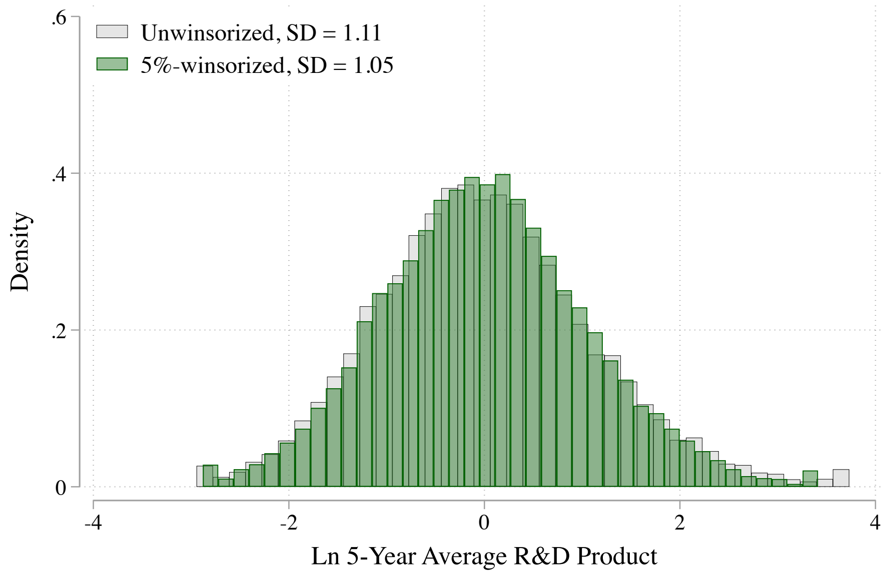
Table B.2: Return Dispersion with Realization Distribution Function

Decay	Return on R&D	Return on Capital		Return on Labor	
Factor p	SD	SD	Δ%	SD	Δ%
100%	1.11	0.77	43.4%	0.79	39.6%
95%	1.09	0.76	43.3%	0.79	37.9%
90%	1.09	0.76	43.3%	0.79	38.2%
75%	1.10	0.76	44.2%	0.79	39.7%
50%	1.14	0.77	48.0%	0.79	43.8%
25%	1.20	0.78	53.7%	0.79	51.1%

Note: Return on R&D assumes decay factor indicated in first column. Alternative return measures are constructed using their original definition, but only taken into account returns when the Return on R&D is non-missing. Column headers SD report standard deviations of return measure. Columns headers Δ% indicate percent difference of Return on R&D dispersion with respect to return in consideration. Returns are measured in logs.

Breakthrough inventions. Another potential source of variation are fat-tailed patent valuations due to rare breakthrough invention. I adjust for this possibility by winsorizing the top 5% of patent valuations within a year and before calculating R&D returns. Figure B.1 confirms that the dispersion remains large with this adjustment.

Figure B.1: Return on R&D Dispersion Adjusted For Breakthrough Inventions



Notes: This figures plots the histogram of the demeaned natural logarithm of the Return on R&D with winsorized and unwinsorized patent valuation. SD in the legend refers to the standard deviation.

Output and input definition. One potential concern with the measured R&D return is that output or input measures might not be comprehensive. On the output side, one might be concerned that patent valuations do not capture all the reward of conducting R&D (Cohen and Klepper, 1996). To address this concern, I follow Bloom et al. (2020) and construct alternative measure of R&D output based on positive changes in revenue, employment, or labor productivity defined as revenue per employee. The alternative measures of the Return on R&D are thus defined as

$$\text{R\&D Return}^X_{it} \equiv \frac{\sum_{s=0}^{4} \max\{\text{X}_{it+s} - X_{it-1+s}, 0\}}{\sum_{s=0}^{4} \text{R\&D Expenditure}_{it-1+s}} \quad \text{with} \quad \text{X} \in \{\text{Revenue, Employment, Labor Productivity}\}. \tag{B.2}$$

As reported in Table B.3, the dispersion using these alternative measures of R&D output turns out to be larger compared to the dispersion in the baseline measure. Thus, Return on R&D dispersion measured using patent valuation is a conservative estimate.

Table B.3: Return Dispersion with Alternative Meausures of R&D Output

Output	Return on R&D	Return on Capital		Return on Labor	
Definition	SD	SD	Δ%	SD	Δ%
Patent valuations	1.11	0.77	43.4%	0.79	39.6%
Revenue changes	1.41	0.77	83.7%	0.79	78.6%
Employment changes	1.71	0.77	122.8%	0.79	117.9%
Labor productivity changes	1.73	0.76	126.6%	0.78	120.9%

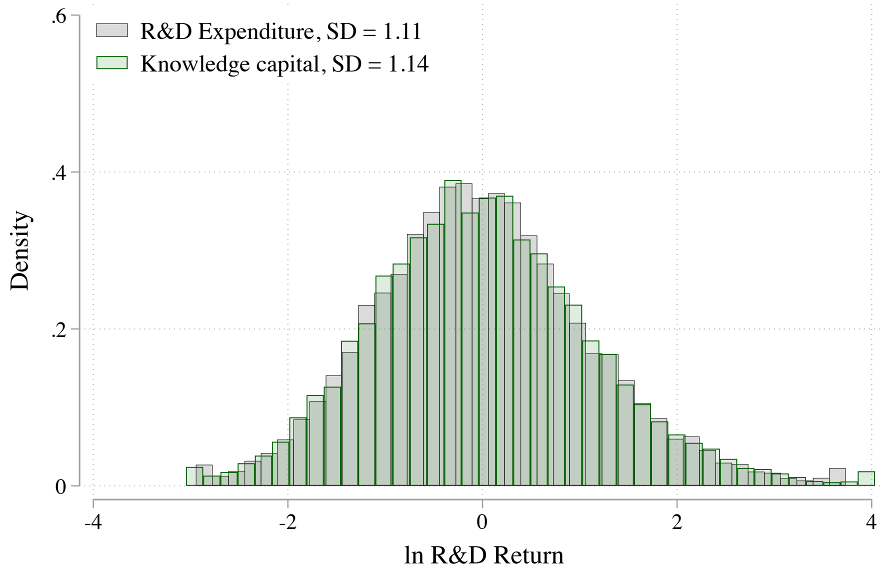
Note: Return on R&D calculated using output definition indicated in first column. Alternative return measures are constructed using their original definition, but only taken into account returns when the Return on R&D is non-missing. Column headers SD report standard deviations of return measure. Columns headers Δ% indicate percent difference of Return on R&D dispersion with respect to return in consideration. Returns are measured in logs.

On the input side, we might be concerned that R&D expenditure does not capture all the inputs associated with the firm's innovation activity. For example, the literature on intangible capital has argued that overhead expenses also serve to enhance a firm's productive capacity, which might be partly reflected in its patent valuation. Building on this insight, I use the knowledge capital series from Ewens et al. (2020), which reflects discounted R&D and overhead expenses, to construct an alternative measure of the Return on R&D as

$$\text{R\&D Return}_{it}^{K} \equiv \frac{\sum_{s=0}^{4} \text{Patent valuations}_{it+s}}{\sum_{s=0}^{4} \text{Knowledge capital}_{it-1+s}}. \tag{B.3}$$

Figure B.2 confirms that the Return on R&D remains highly dispersed using the alternative input measure. In fact, the dispersion increases slightly from 1.10 to 1.15.

Figure B.2: Return on R&D Dispersion With Alternative Input Measures



Notes: This figures plots the histogram of the demeaned natural logarithm of the Return on R&D under alternative input specifications. SD in the legend refers to the standard deviation.

Merger & Acquisitions. The acquisition of innovative startups by established firms is a common phenomenon, which has received increasing attention in the business press and academic literature in recent years (Cunningham et al., 2020). One potential concern in my context is that some patents were developed by an acquisition target prior to the grant date and subsequently attributed to the acquirer, while the associated R&D expenditure is not recorded by the acquirer. This sequence of events induces measurement error in the return on R&D as valuations are recorded without the associated expenditure.

I investigate this possibility using two complementary approaches. First, I regress the log returns on the ratio of acquisitions to R&D expenditure. As long as the under-counted R&D expenditure is small relative to total R&D expenditure, this identifies the average share of acquisitions attributable as R&D. Formally, let $\mathtt{V}$ be the value created from innovation, $\mathtt{R\&D}$ the recorded R&D expenditure, $\mathtt{AQC}$ the total value of acquisitions, and s the share thereof attributable to R&D expenditure. In the standard model we then have

$$\frac{\mathtt{V}}{\mathtt{R\&D}} = (1+\phi) \times \left(1 + s\frac{\mathtt{AQC}}{\mathtt{R\&D}}\right),$$

where $1+\phi$ is the true return on R&D. See Section 4.

As long as $s\frac{\texttt{AQC}}{\texttt{R\&D}}$ is close to zero, we thus have $\partial \ln \frac{\texttt{V}}{\texttt{R\&D}} / \partial \frac{\texttt{AQC}}{\texttt{R\&D}} \approx s$. Thus, regressing log returns on the valuation to R&D share identifies the average share of acquisition expenditure attributable to R&D expenditure.

Column (1) in Table B.4 estimates this relationships and finds an average s of 5.5% suggesting that acqusitions might indeed be linked to measured returns on R&D. When calculating the adjusted returns, however, I find that they are more, rather than less dispersed with a standard deviation of 1.13 compared to the baseline of 1.11.

Second, the sequence of events described above suggests that the inventors listed on the patent should not be linked previously to the acquiring firm. We can thus get an insight in the potential prevalence of this phenomenon by looking at the share of valuations linked to inventor teams that have not previously worked at the firm. Under the hypothesis developed above, we would expect that firms with a large share of valuations attributable to new teams should have higher returns on R&D as we under-count R&D expenditure.

Columns (2)-(5) in Table B.4 show that this is not the case. Across alternative definitions of new team, I find that firms with a large share of valuations accounted for by new teams have lower, not higher, R&D returns.

Together, the evidence thus suggest that under-counting of R&D due to acquisitions is not a quantitatively large source of measured return on R&D variation.

Table B.4: Returns on R&D and New Inventor Teams

	(1)	(2)	(3)	(4)	(5)
			ln **R&D Return**		
ln Acquisition / R&D	0.055***				
	(0.007)				
New Team Share		-0.855*			
		(0.491)			
NTS (Experienced)			-0.627***		
			(0.223)		
Majority NTS				-0.606	
				(0.377)	
Majority NTS (Experienced)					-0.535***
					(0.205)
Within R2	0.03	0.00	0.01	0.00	0.01
Observations	11,845	11,845	11,845	11,845	11,845

Note: This table reports OLS coefficient estimates. New Team Share (NTS) is the fraction of patent valuation attributable to inventor teams in which none of the team members has previously been associated with the firm. Majority New Team Share is the equivalent concept requiring the majority of inventors being previously unconnected to the firm. Experienced versions focus on inventors with previous patents. Regressions control for NAICS3 $\times$ Year effects and cluster standard errors at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

B.2 Measurement Error

One potential source of variation in the measured Expected Return on R&D is measurement error. In this section I propose two complementary approaches to estimating the contribution of measurement error to measured Expected Return on R&D dispersion. I begin by taking a structural approach using a GMM estimator to investigate the importance of classical measurement error. In addition, I use bootstrapping to estimating the potential measurement error due the uncertainty around patent valuations.

GMM Estimation of Measurement Error. Consider a stationary, AR(1) process $\{y_{it}\}$:

$$y_{it} = (1-\rho)\mu_i + \rho y_{it-1} + \varepsilon_{it} \quad \text{with} \quad \varepsilon_{it} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \quad \text{and} \quad \mu_i \sim N(0, \sigma_\mu^2). \tag{B.4}$$

The econometrician observes the process with i.i.d. normal measurement error:

$$\tilde{y}_{it} \equiv y_{it} + \nu_{it} \quad \nu_{it} \overset{iid}{\sim} N(0, \sigma_\nu^2). \tag{B.5}$$

Lemma B.1. *Define $\Delta\tilde{y}_{it} \equiv \tilde{y}_{it} - \tilde{y}_{it-1}$, then under $\rho \in (0,1)$, we have*

$$\begin{aligned}
m_1 &\equiv Cov(\tilde{y}_{i,t}, \Delta\tilde{y}_{it}) = \frac{1}{1+\rho}\sigma_\varepsilon^2 + \sigma_\nu^2 \\
m_2 &\equiv Cov(\tilde{y}_{i,t}, \Delta\tilde{y}_{it-1}) = \frac{\rho}{1+\rho}\sigma_\varepsilon^2 \\
m_3 &\equiv Cov(\tilde{y}_{i,t}, \Delta\tilde{y}_{it-2}) = \frac{\rho^2}{1+\rho}\sigma_\varepsilon^2 \\
m_4 &\equiv Cov(\tilde{y}_{i,t}, \tilde{y}_{it-1}) = \sigma_\mu^2 + \frac{\rho}{1-\rho^2}\sigma_\varepsilon^2.
\end{aligned}$$

Proposition B.1. *If $\rho \in (0,1)$, we can solve for $\{\rho, \sigma_\mu, \sigma_\varepsilon, \sigma_\nu\}$ using the population auto-covariance structure of $\tilde{y}_{it}$ and $\Delta\tilde{y}_{it} \equiv y_{it} - y_{it-1}$:*

$$\beta \equiv \begin{bmatrix} \rho \\ \sigma_\varepsilon^2 \\ \sigma_\mu^2 \\ \sigma_\nu^2 \end{bmatrix} = \begin{bmatrix} \frac{m_3}{m_2} \\ \frac{(m_2)^2}{m_3} + m_2 \\ m_4 - \frac{(m_2)^2}{m_2 - m_3} \\ m_1 - \frac{(m_2)^2}{m_3} \end{bmatrix}$$

Let Ω be the covariance matrix of m and denote the sample moments by $\hat{m}$, then

$$\hat{\beta} \sim N(\beta, \Sigma) \quad \textit{and a feasible estimator is} \quad \hat{\Sigma} = \left(\frac{\partial\hat{\beta}}{\partial m}\right)' \hat{\Omega} \left(\frac{\partial\hat{\beta}}{\partial m}\right),$$

where $\partial\beta/\partial m$ is evaluated at $\hat{m}$ and given by

$$\frac{\partial\beta}{\partial m} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ -\frac{m_3}{(m_2)^2} & 2\frac{m_2}{m_3} + 1 & m_2\left(\frac{m_2 - 2m_3}{(m_2-m_3)^2}\right) & -2\frac{m_2}{m_3} \\ \frac{1}{m_2} & -\left(\frac{m_2}{m_3}\right)^2 & -\left(\frac{m_2}{m_2-m_3}\right)^2 & -\left(\frac{m_2}{m_3}\right)^2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Proof. The first part follows by rearranging the moments expressions. The second part follows from the Law of Large Numbers for the moment vector and the Delta method. $\square$

Note that this methodology does not aggregate. In particular, if we assume that Expected Return on R&D follows an AR(1) in logs at the annual level, we cannot implement the above methodology at the 5-year horizon directly as the 5-year Expected Return on R&D is a weighted-average of the annual Return in levels, which does not translate into logs:

$$\frac{\sum_{s=0}^{4} \text{Pat. Val.}_{it+s}}{\sum_{s=0} \text{R\&D Exp.}_{it-1+s}} = \sum_{s=0}^{4} \frac{\text{R\&D Exp.}_{it-1+s}}{\sum_{w=0} \text{R\&D Exp.}_{it-1+w}} \times \frac{\text{Pat. Val.}_{it+s}}{\text{R\&D Exp.}_{it-1+s}}.$$

To address this concern, I will estimate the system at the 1-year level and propose a methodology to estimate the importance of measurement error at the 5-year level. I restrict my sample to 1-year returns with at least 5 patents and provide additional estimates for an adjusted measure with at least 10 patents per return, winsorized patent valuations, and only within industry-year variation.

The GMM estimates presented in Table B.5 suggest that measurement error constitutes little of the overall variation in the 1-year Return on R&D. The estimated measurement error variation is around 0.04 for both return measures, but only significantly different from 0 at the 5% level for the adjusted returns. In addition, I consistently find that the Return on R&D is highly auto-correlated with significant variation due to idiosyncratic shocks. The results for permanent differences across firms are mixed. While the baseline estimates suggest some role for them, the estimated coefficient for the adjusted returns is very close to 0. One interpretation is that there are permanent differences across industries, which do not show up for the adjusted returns as they are residualized. Note, however, that the standard errors around the estimates for σ_μ^2 are very large.

Table B.5: GMM results for AR(1) with Noise

Parameter	R&D Return	
	Baseline	Within
ρ	0.860***	0.892***
	(0.035)	(0.062)
σ^2_ε	0.287***	0.170***
	(0.025)	(0.020)
σ^2_μ	0.292	-0.046
	(0.976)	(0.861)
σ^2_ν	0.041	0.044**
	(0.035)	(0.017)
Observations	10,424	7,553

Note: Standard errors clustered at the NAICS6 level and reported in brackets.

As discussed before, we cannot immediately translate these estimates into measurement error contributions at the 5-year level due to aggregation. I address this challenge by adding some structure on the firm R&D process. In particular, I will assume that each firm in my data solves the simple maximization problem

$$\max_{\ell_{it}} \left\{ \varphi \ell_{it}^{\frac{1}{1+\phi}} - \Delta_{it} \times W \ell_{it} \right\}. \tag{B.6}$$

The source of R&D returns in this framework is Δ_{it} and I will consequently assume that it follows an AR(1) process, which the researcher observed with i.i.d. measurement error.

Lemma B.2. *Under above assumptions, the 5-year Return on R&D is given by*

$$\textit{Expected Return on R\&D}_{it} = (1+\phi) \times \frac{\sum_{s=0}^{4} \Delta_{it}^{-\frac{1+\phi}{\phi}} \times \tilde{\Delta}_{it}}{\sum_{s=0}^{4} \Delta_{it}^{-\frac{1+\phi}{\phi}}}.$$

Proof. The solution to the firm optimization problem is given by

$$\ell_{it} = (\Delta_{it}(1+\phi)W)^{-\frac{1+\phi}{\phi}} \times (\varphi)^{\frac{1+\phi}{\phi}}$$

The annual return on R&D is proportional to Δ_{it}:

$$\frac{\varphi \ell_{it}^{\frac{1}{1+\phi}}}{W\ell_{it}} = (1+\phi) \times \Delta_{it}.$$

By definition, we can then express the overall return as measured in the data as

$$\frac{\sum_{s=0}^{4} W\ell_{it+s} \times \tilde{\Delta}_{it+s}}{\sum_{s=0}^{4} W\ell_{it+s}} = (1+\phi) \times \frac{\sum_{s=0}^{4} \Delta_{it}^{-\frac{1+\phi}{\phi}} \times \tilde{\Delta}_{it}}{\sum_{s=0}^{4} \Delta_{it}^{-\frac{1+\phi}{\phi}}}.$$

□

Using this framework, we can simulate data based on the estimates in B.5 and aggregate to the 5-year level as suggested above. To estimate the importance of measurement error, we can then compare baseline estimates against a counterfactual with $\sigma_\nu^2 = 0$. I follow the literature and set $\phi = 1$ for the purpose of this exercise Acemoglu et al. (2018).

Table B.6 reports the results, which suggest that measurement error makes a minor contribution to the dispersion in the Expected Return on R&D. For baseline and adjusted returns I find that measurement error contributes less than 1% to the overall dispersion in the 5-year Expected Return on R&D. The importance of measurement error is decreasing in the time-horizon considered as the individual shocks average out.

Table B.6: Disperion in Simulated Expected Return on R&D

Measure	Baseline		Within	
	1-year	5-year	1-year	5-year
SD	1.191	1.099	0.937	0.854
SD with $\sigma_\nu^2 = 0$	1.174	1.094	0.913	0.848
$\Delta\%$	1.5%	0.4%	2.5%	0.7%

Note: The first data row reports the standard deviation of the simulated Expected Return on R&D using the associated GMM parameter estimates. The second row recalculates this dispersion imposing no measurement error or $\sigma_\nu^2 = 0$. The final row reports the reduction in the dispersion of the Expected Return on R&D due to the reduction in measurement error.

Bootstrap Estimation for Valuation Uncertainty. In addition to the investigation of classical measurement error, I consider the role of patent valuation uncertainty explicitly. In a bootstrap procedure I redraw patent valuations from the realized patent portfolio and construct Returns on R&D assuming that the first targets a return proportional the expected value of patent valuations ex-ante. Repeating this exercise for 1000 iteration I then calculate an estimated dispersion in the measured Expected Return on R&D based on uncertain valuation outcomes only.

Each iteration in my procedure proceeds as follows:

1. For each firm and 5-year window in which the firm has at least 50 patents:
   (a) From the portfolio of patent valuations for the firm-period, draw with replacement an alternative portfolio with as many valuations as the firm had patents in the period.
   (b) Calculate the return as the ratio the valuations in the alternative portfolio divided by the valuation of the true portfolio.
2. Calculate the standard deviation of Return on R&D for the simulated data.

I repeat this procedure until I have 1000 bootstrap estimates of the standard deviation of Returns on R&D. Note that the resulting dispersion in the Return on R&D is driven exclusively by the variability of patent valuations and would yield 0 variation if all patents had the same value.

One way to interpret this approach is that the realized patent portfolio is a good approximation for the true uncertainty faced by the firm around its innovation outcomes. The procedure ignores all variation coming from shifts in the level of expected patent valuation and instead considers the dispersion conditional on the average value only. As a result, the procedure will overstate the associated measurement error if firms are aware that certain project are low or high expected value within their research portfolio.

Table B.7 reports estimates suggesting that the measurement error due to patent valuation uncertainty could account for up to $0.067/1.1 = 6\%$ of the standard deviation of the measured Expected Return on R&D. Unsurprisingly, the estimated measurement error declines with the size of the minimum patent portfolio and is precisely estimated with tight confidence intervals.

Table B.7: Measurement Error Estimates using Bootstrap Procedure

Minimum patents	Estimate	Standard error	95% Confidence Interval
30	0.08	(0.002)	[0.077,0.084]
50	0.067	(0.002)	[0.064,0.071]
100	0.051	(0.001)	[0.049,0.053]
200	0.042	(0.001)	[0.04,0.045]

Note: Measurement error estimates based on distribution of patent valuations for different cut-offs levels of minimum patent counts.

B.3 Additional Evidence on Systematic Drivers

Risk. I investigate the importance of risk for R&D returns in an OLS framework using four risk proxies. Firstly, I use the CAPM β as provided by WRDS, where I map firms to securities using the same crosswalk as for the mapping between the patent data and Compustat. Secondly, I estimate the long-run firm β using an OLS regression of firm-specific excess returns on market excess returns at the annual level. All return data is from WRDS. Finally, I calculate innovation specific βs using the same framework by calculating R&D returns at the 1- and 5-year level for firms with at least 5 patents over the horizon and regressing them on market excess returns.

To estimate the long-run firm β, I first calculate the annual stock market return for the firm and the S&P500 index. I subtract the risk free rate from both to construct excess returns and regress the firm-specific excess return on the market excess return firm-by-firm to construct firm-level stock market βs. For the innovation-based measures I follow a similar approach, but replace firms' stock market return with the Return on R&D. I calculate this measure for both the 1-year and 5-year Return on R&D, where I restrict both to observations with at least 5 patents. I then regress the innovation-based excess returns on the market return firm-by-firm to construct firm-specific, innovation-based risk factors $\beta_{R\&D}$.

Table B.8 reports the OLS regression results relating the risk measures to the Expected Return on R&D. As discussed in the main text, I find no correlation with the general, stock return-based risk measures, but significant correlations with the innovation specific risk factors.

Table B.8: Return on R&D and Firm-level Risk

	(1)	(2)	(3)	(4)
		Return on R&D		
Compustat β_{CAPM}	0.005			
	(0.060)			
$\hat{\beta}_{CAPM}$		-0.041		
		(0.078)		
1-year $\hat{\beta}_{R\&D}$			0.021***	
			(0.003)	
5-year $\hat{\beta}_{R\&D}$				0.015***
				(0.005)
R2	0.40	0.30	0.35	0.32
Within R2	0.00	0.00	0.08	0.02
Observations	6,797	10,164	9,587	9,858

Note: All regressions control for NAICS3 × year fixed effects. All returns are in logs. Standard errors clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Labor Market Dominance. Labor market dominance has been closely connected with labor market power (Berger et al., 2022; Yeh et al., 2022). Furthermore, dominance has the added feature that it connects labor market power with firm size. I construct a measure of labor market dominance in the market for inventors to investigate the potential connection between dominance and R&D returns. For each new patent in a firm's portfolio I calculate the share of potential inventors that are working with the firm, where I classify someone as a potential inventor if they work on patents with the identical technology classification. I then average this measure out over all of the firm's patent to get a measure of overall inventor market dominance. See Appendix A for further details on the construction.

Column (1) in Table B.9 reports the OLS coefficient of a regression of the R&D return on the dominance measure. In line with a monopsony interpretation, I find that dominant firms have higher returns. A one standard deviation higher dominance measure is associated with 14% larger return.[32] A potentially confounding factor is firm size, which could be linked to returns through alternative mechanisms. Column (2) confirms that the link between dominance and returns remains strong even when controlling for inventor employment.

Table B.9: Return on R&D, Labor Market Dominance, and Specialization

	(1)	(2)	(3)	(4)
	ln **Return on R&D**			
ln Dominance	0.140***	0.098**		
	(0.041)	(0.040)		
ln Specialization			0.300***	0.282***
			(0.090)	(0.087)
ln Inventors		0.222***		0.221***
		(0.035)		(0.032)
R2	0.01	0.08	0.01	0.07
Observations	10,444	10,444	11,795	11,795

Note: This table reports OLS regression coefficients. See Appendix A for variable definitions. Standard errors are clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Inventor Specialization. Inventor specialization is another potential source of employer bargaining power as it reduces the set of potential employers. I investigate its relationship with R&D returns by aggregating inventor-level specialization measures to the firm-level. For an individual inventor, I construct a specialization measure based on the cosine distance

[32] The standard deviation of ln Dominance is 1.01 s.t. $1.01 \times 0.14 \approx 0.14$. In turn, $\exp(0.14) - 1 \approx 14\%$.

between the technology classifications of patents that the inventor worked on over the period. I then average this measure to the firm-level by taking a patent-weighted average over inventors associated with the firm. See Appendix A for further details on the construction.

Column (3) in Table B.9 reports the OLS coefficient of a regression of the R&D return on the specialization measure. Indeed, I find that firms with more specialized inventors have higher returns on R&D, which supports a labor market power interpretation. A one standard deviation larger specialization measure is associated with an 8% larger return on R&D. As shown in column (4), this relationship is not driven by firm-size differences.

State and university partnerships. State involvement and university partnerships can create measured R&D return dispersion unconnected to economic fundamentals if they lead to inaccurate measurement of R&D inputs and outputs. For example, R&D subsidies reduce the effective cost of R&D to the firm, which is not reflected in gross R&D expenditure as reported in the firms accounting statements. Similarly, suppose the firm engage in a research partnership with a university with the agreement that all patents are assigned to the firm. Again, this scenario could lead us to under-count the true cost of R&D associated with inventions or alternatively overstate the value created by the firm's own R&D expenditure. I will consider two complementary approaches to adjusted for these measurement concerns.

First, I directly identify inventions created with state support or university partnerships using patent records. I classify assignees into governmental institutions, universities, or neither based on the listed name using key words such as "university" or "federal agency". Furthermore, I classify patents as government related if they have a public interest statement, which indicates that a federal agency has supported the invention and/or has remaining rights over the patent. Following this procedure, I can then classify whether a patent is related to a government agency, university, or neither, and calculate their value at the firm level.

Table B.10 confirms that the dispersion is essentially unaffected by adjusting for the valuations associated with university and state involvement. Excluding them either partly or full does not affect measured dispersed up to the second digit, which is driven by the fact that only 1.2% of valuations in the sample are linked state or university involvement, whereof 1/3 is linked to universities and 2/3 to state involvement. It appears that these collaborations are not quantitatively important for patent valuations created by firms in my sample nor the documented R&D return dispersion.

Table B.10: Return on R&D Dispersion with Valuation Adjusted for State and University Collaborations

Actor Considered	Unadjusted SD	50% Adjusted SD	50% Adjusted Δ%	100% Adjusted SD	100% Adjusted Δ%
University	1.11	1.11	0.0%	1.11	0.1%
State	1.11	1.11	-0.0%	1.11	0.0%
Both	1.11	1.11	0.0%	1.11	0.1%

Note: Return on R&D calculated using adjusted valuations. First column indicates the adjusted patent types. 50% and 100% adjustment refer to subtracting the respective percentage of the patent valuations from collaborations from the total valuations. Column headers SD report standard deviations of return measure. Columns headers Δ% indicate percent difference of Return on R&D dispersion with respect to unadjusted return. Returns are measured in logs.

Secondly, I directly adjust R&D spending for state-level R&D subsidies using the data in Lucking (2019), where I map firms to states either via their headquarter location or the distribution of inventors associated with the firms' patents as recorded in Berkes (2016). I then calculate the adjusted R&D spending as R&D expenditure times user cost. I experiment with full or partial adjustments to account for the fact that state-level subsidies typically apply to the marginal R&D expenditure. Table B.11 confirms that user cost differences are not a driver of Return on R&D dispersion. On the contrary, taking into account the full R&D user cost increases the dispersion by about 1%.

Table B.11: Return on R&D Dispersion with Valuation Adjusted for State R&D Tax Credits

Geography Matching	Unadjusted SD	50% Adjusted SD	50% Adjusted Δ%	100% Adjusted SD	100% Adjusted Δ%
Headquarters	1.11	1.11	0.1%	1.12	1.3%
Innovators	1.11	1.11	0.0%	1.12	1.0%

Note: Return on R&D calculated using adjusted valuations. First column indicates the adjusted patent types. 50% and 100% adjustment refer to subtracting the respective percentage of the patent valuations from collaborations from the total valuations. Column headers SD report standard deviations of return measure. Columns headers Δ% indicate percent difference of Return on R&D dispersion with respect to unadjusted return. Returns are measured in logs.

B.4 Labor Supply Elasticity Estimates

Table B.12 reports the first-stage results for the main specification. In addition I report results for the specification controlling for lagged wage and employment growth in Table B.14. This specification is similar to the main specification in Seegmiller (2021).

Table B.12: Inventor Inverse Labor Elasticity Estimates — First Stage

	(1)	(2)	(3)	(4)
A. Main		$\Delta \ln \texttt{Inventors}_{it}$		
Stock Return$_{it}$	0.066***	0.043***	0.044***	0.035***
	(0.007)	(0.009)	(0.006)	(0.009)
— × {Top 50% R&D Return$_{it}$}		0.043***		0.016
		(0.011)		(0.010)
B. Interaction		$\Delta \ln \texttt{Inventors}_{it}$ × **{Top 50% R&D Return$_{it}$}**		
Stock Return$_{it}$		0.002		-0.002
		(0.002)		(0.002)
— × {Top 50% R&D Return$_{it}$}		0.047***		0.033***
		(0.007)		(0.005)
Firm Effects			✓	✓
First stage F stat. (Main)		41		31
First stage F stat. (Inter.)		41		31
Observations	14,907	14,907	14,890	14,890

Note: Controls include lagged inventor wage and employment growth as well as current inventor productivity growth. All regressions control for NAICS3 × year fixed effects. Standard errors clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.13: Inventor Inverse Labor Elasticity Estimates

	(1)	(2)	(3)	(4)
		$\Delta \ln \texttt{Inventor Wage}_{it}$		
$\Delta \ln \texttt{Inventors}_{it}$	0.923***	0.756**	1.449***	1.211***
	(0.195)	(0.313)	(0.370)	(0.422)
— × {Top 50% R&D Return$_{it}$}		1.119**		2.015***
		(0.496)		(0.767)
{Top 50% R&D Return$_{it}$}		-0.219***		-0.358***
		(0.043)		(0.071)
Firm Effects			✓	✓
First stage F stat. (Main)	101	41	46	31
First stage F stat. (Inter.)		61		45
Observations	14,907	14,907	14,890	14,890

Note: Controls include lagged inventor wage and employment growth as well as current inventor productivity growth. All regressions control for NAICS3 × year fixed effects. Standard errors clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.14: Inventor Inverse Labor Elasticity Estimates With Controls

	(1)	(2)	(3)	(4)
A. Second stage		$\Delta \ln$ **`Inventor Wage`**$_{it}$		
$\Delta \ln$ `Inventors`$_{it}$	4.826***	3.818***	4.570***	3.620***
	(0.980)	(0.981)	(1.045)	(0.931)
— $\times$ {Top 50% R&D Return$_{it}$}		2.352***		2.950**
		(0.816)		(1.228)
{Top 50% R&D Return$_{it}$}		-0.142***		-0.201***
		(0.050)		(0.063)
B. First Stage: Main		$\Delta \ln$ **`Inventors`**$_{it}$		
Stock Return$_{it}$	0.066***	0.023***	0.022***	0.025***
	(0.006)	(0.005)	(0.004)	(0.006)
— $\times$ {Top 50% R&D Return$_{it}$}		0.002		-0.006
		(0.006)		(0.007)
C. First Stage: Interaction		$\Delta \ln$ **`Inventors`**$_{it}$ $\times$ **{Top 50% R&D Return$_{it}$}**		
Stock Return$_{it}$		-0.005		-0.004
		(0.003)		(0.002)
— $\times$ {Top 50% R&D Return$_{it}$}		0.034***		0.027***
		(0.007)		(0.005)
Firm Effects			✓	✓
First stage F stat. (Main)		37		32
First stage F stat. (Inter.)		37		32
Observations	14,044	14,044	14,028	14,028

Note: All regression control for lagged inventor wage and employment growth as well as current inventor productivity growth. All regressions control for NAICS3 $\times$ year fixed effects. Standard errors clustered at the NAICS6 level.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

C Model Appendix

C.1 Proofs and Further Results for Section 4

Proof of Lemma 1. Under the stated assumptions, $\varepsilon_C = 1$ and $\varepsilon_F = \frac{1}{1+\phi}$. The result follows from (4). $\square$

Proof of Lemma 2. This follows immediately from the production and cost function together with the first order condition. $\square$

Lemma C.1. *Consider a firm solving (3) s.t. to financial frictions putting a constraint on its maximal investment level: $\ell \leq \bar{\ell}$. Furthermore, the firm borrows to finance its R&D expenditure at net interest rate r. Then, its return on R&D is given by*

$$\frac{F(\ell^*)V}{C(\ell^*)} = \frac{\varepsilon_C}{\varepsilon_F} \times (1 + \lambda(\ell^*)) \times (1 + r) \tag{C.1}$$

where $\lambda(\ell^*)$ measured to degree to which the borrowing constraint is binding with $\lambda(\ell^*) = 0$ for $\ell^* < \bar{\ell}$ and $\lambda(\ell^*) > 0$ otherwise.

Proof of Lemma C.1. The firm's problem is given by

$$\max_{\ell} \left\{ F(\ell)V - (1+r)C(\ell), \text{ s.t. } \ell \leq \bar{\ell} \right\}$$

The first order condition yield equation (C.1), where $\lambda(\ell^*)$ is the Lagrange multiplier on the capacity constraint. The multiplier is positive if the constraint binds and zero otherwise. $\square$

Lemma C.2. *Consider a firm solving (3) s.t. R&D cost function $C(\ell) = \mathcal{W}(\ell) \cdot \ell$, where $W(\ell)$ is the firm-specific wage. Then, its return on R&D is given by*

$$\frac{F(\ell^*)V}{C(\ell^*)} = \frac{1}{\varepsilon_F} \times (1 + \varepsilon_W) \quad \textit{with} \quad \varepsilon_W = \frac{\partial \ln W(\ell)}{\partial \ln \ell}\Big|_{\ell=\ell^*}. \tag{C.2}$$

In contrast, if the firm ignores its impact on wages, the return on R&D is given by

$$\frac{F(\ell^*)V}{C(\ell^*)} = \frac{1}{\varepsilon_F}. \tag{C.3}$$

Firms taking into account their impact on wages have lower input demand ℓ^*.

Proof of Lemma C.2. The proof follows immediately from the first order conditions. $\square$

C.2 Proofs and Further Results for Section 5

Definition 3. *A Balanced Growth Path equilibrium is a competitive equilibrium such that prices W_t and quantities $\{Y_t, A_t\}$ grow at a constant rate g and R_t is a constant.*

As summarized in Proposition C.1, the model formulation allows me to decompose the value function into a profit and R&D component. The former captures the expected net-present-value of the profits associated with existing leadership positions. The latter capture the value of the firm's ability to conduct R&D and create leadership positions in the future. The assumption delivering this property is the fixed R&D production function.

Proposition C.1. *Along the BGP, the normalized value function $V(\cdot) \equiv V_t(\cdot)/Y_t$ is constant and can be decomposed into a profit and R&D component:*

$$V(\cdot) = \sum_{j \in \mathcal{J}_{it}} \mathcal{V}(\lambda_{jt}) + \tilde{V}(\ell_{it-1}, \varphi_{it}) \tag{C.4}$$

The profit component is equivalent to the expected discounted sum of profits:

$$\mathcal{V}(\lambda_{jt}) = \frac{\pi(\lambda_{jt})}{1 - \beta(1-z)} \qquad \textit{with} \qquad \pi(\lambda_{jt}) = 1 - 1/\lambda_{jt}. \tag{C.5}$$

The R&D component is the solution to the value function maximization problem

$$\tilde{V}(\ell_{it-1}, \varphi_{it}) = \max_{\ell_{it}} \left\{ -C(\ell_{it-1}, \ell_{it}) + \beta \left(z_{it} \mathbb{E}_\lambda[\mathcal{V}(\lambda)] + \mathbb{E}_t[\tilde{V}(\ell_{it}, \varphi_{it+1})] \right) \right\}, \tag{C.6}$$

where expectations $\mathbb{E}_t[\cdot]$ are taken with respect to the productivity process only and expectation $\mathbb{E}_\lambda[\cdot]$ capture the distribution over λ.

Proof of Proposition C.1. Firstly, we can guess and verify that the value function is proportional to Y_t, since profits are proportional to Y_t and cost are proportional to W_t with $W \equiv W_t/Y_t$ being constant along the balanced growth path by assumption. The Euler equation then implies $\frac{1+g}{R} = \beta$ and we have

$$V(\mathcal{A}_{it}, \varphi_{it}, \ell_{it-1}) = \max_{\ell_{it}} \Big\{ \sum_{j \in \mathcal{J}_{it}} \pi(\lambda_{jt}) - C(\ell_{it-1}, \ell_{it}) + \beta \mathbb{E}_t \big[V(\mathcal{A}_{it+1}, \varphi_{it+1}, \ell_{it}) \big] \Big\}, \tag{C.7}$$

where $C(\ell_{it-1}, \ell_{it}) \equiv C_t(\ell_{it-1}, \ell_{it})/Y_t$.

Secondly, we can guess and verify

$$
\begin{aligned}
&V(\mathcal{A}_{it}, \varphi_{it}, \ell_{it-1}) = \tilde{V}(\varphi_{it}, \ell_{it-1}) + \sum_{j \in \mathcal{J}_{it}} \mathcal{V}(\lambda_{jt}) \qquad \text{with} \\
&\tilde{V}(\ell_{it-1}, \varphi_{it}) = \max_{\ell_{it}} \left\{ -C(\ell_{it-1}, \ell_{it}) + \beta \left(z_{it} \mathbb{E}_{\lambda}[\mathcal{V}(\lambda)] + \mathbb{E}_t[\tilde{V}(\ell_{it}, \varphi_{it+1})] \right) \right\} \\
&\mathcal{V}(\lambda_{jt}) = \pi(\lambda_{jt}) + \beta(1 - z_t)\mathcal{V}(\lambda_{jt}).
\end{aligned}
\tag{C.8}
$$

The intuition behind this form is that innovation and product market activity do not interact from the perspective of the firm and are thus separable from the perspective of the firm. Furthermore, the firms product lines do not interact with each other and, thus, again are separable. □

Lemma C.3. *The growth rate in the economy is given by*

$$
g = \int_0^1 z_{it}(\mathbb{E}[\lambda] - 1)di = z \cdot \mathbb{E}[\ln \lambda] \tag{C.9}
$$

where $z \equiv \int_0^1 z_{it} di$ is the aggregate innovation rate, which is constant along the BGP.

Proof of Lemma C.3.

$$
\begin{aligned}
g = \frac{A_{t+1} - A_t}{A_t} &\approx \ln\left(A_{t+1}/A_t\right) = \int_0^1 \ln\left(A_{jt+1}/A_{jt}\right) dj \\
&= \int_0^1 \left(z_{it} \ln \lambda_{it} + (1 - z_{it}) \ln(1)\right) di = \left(\int_0^1 z_{it} di\right) \mathbb{E}[\ln \lambda_{it}]
\end{aligned}
$$

The approximation holds for low values of g, which is applicable in this case. The second equality simply introduces the definition of A_t. The first equality in the second line follows as each product line has the same probability to be innovated on my a random firm such that the expected improvement is simply the expected improvement made by a random firm. A random firm improves upon a product line by λ_{it}, which is not known in advance, with probability z_{it} and makes no improvement otherwise. Finally, since the improvement size it not known in advance, $\ln \lambda_{it}$ is independent of z_{it}, which leads to the second equality in the second line. □

C.3 Microfounding the Wage Function in GE

The wage function can be motivated in a model where R&D workers have a preference for an even distribution across firms. I will present the static problem below, which can be embedded in infinite horizon straight-forwardly.

$$\begin{aligned} &\max_{\{\ell_{it},c_t\}} \quad \log c_t - \frac{\alpha}{1+\xi} \times \left(\int_0^1 \left(\frac{\ell_{it}}{L} \right)^{1+\xi} di \right) \\ &\text{s.t.} \quad c_t = \int_0^1 \left(\frac{\ell_{it}}{L} \right) W_{it} di \quad \text{and} \quad \int_0^1 \ell_{it} di = L \end{aligned} \tag{C.10}$$

There is a unit mass of inventors, which are hand to mouth in equilibrium. Inventors like consumption, which they finance through wage income. Workers face different wages by across firms, which they take as given, and optimize allocation across firms subject to a disutility for dispersed allocations.

The household problem gives rise to the simple wage function

$$W_{it} = W_t \left(1 + \alpha \left(\left(\frac{\ell_{it}}{L} \right)^{\xi} - \int_0^1 \left(\frac{\ell_{it}}{L} \right)^{1+\xi} di \right) \right). \tag{C.11}$$

Firms with relatively high wages receive relatively more workers, but the strength of this mechanism depends on α and ξ. The wage elasticity is given by

$$\frac{\partial \ln W_{it}}{\partial \ln \ell_{it}} = \xi \times \frac{\left(\frac{\ell_{it}}{L}\right)^{\xi}}{\frac{1}{\alpha} - \int_0^1 \left(\frac{\ell_{it}}{L}\right)^{1+\xi} di + \left(\frac{\ell_{it}}{L}\right)^{\xi}}, \tag{C.12}$$

Note that this formulation coincides with the wage function in the main text for $\bar{\ell} = \frac{1}{\alpha} - \int_0^1 \left(\frac{\ell_{it}}{L}\right)^{1+\xi} di$. In this formulation, a lower α performs a similar role to a large $\bar{\ell}$, by flattening the elasticity profile. The formulation in the text reduces the computational burden of solving the model.

C.4 Numerical Solution

I employ two solution algorithms when solving the model, one for moment matching and one for counterfactuals. The first one imposes a growth rate of 1.5% exogenously and sets the average R&D productivity μ accordingly. The second one takes R&D productivity as given and solves for the growth rate of the economy.

Algorithm with exogenous growth rate. The model with fixed growth rate has two features that I will take advantage of. Firstly, I can solve for the equilibrium innovation rate directly using the definition of the growth rate one I've fixed the process for λ

$$z = \frac{g}{\mathbb{E}[\ln \lambda]}.$$

This is useful as it pins down the equilibrium discount rate in the economy without requiring another loop.

Secondly, wages are directly proportional to the aggregate R&D productivity level. Thus, once we have solved for the equilibrium allocation of labor across firms for an arbitrary productivity level with a market clearing wage, we can scale the wage and R&D productivity such that the allocation remains the same, but the economy achieves the required innovation rate z.

My algorithm then involves an inner and an outer loop. In the inner loop I solve for firms' optimal R&D policy and the resulting steady state distribution using standard value function iteration with Howard improvement steps and non-stochastic simulation. In the outer loop, I use a bisection algorithm to determine the equilibrium wage that clears the inventor market for a given average productivity level.

Once the outer loop converged, I calculate the innovation rate under the average R&D productivity level and then scale wages and R&D productivity to achieve the required innovation rate to achieve a growth rate of 1.5% per year. I confirm my guess by solving the model under this parameterization and calculating the model's growth rate.

Algorithm with endogenous growth rate. In the algorithm with endogenous growth rate I instead fix the average R&D level and proceed in three loops. The inner two loops are described above and solve for firms' optimal policy, allocation across states, and R&D wage for a given innovation rate. In the outer loop I then solve for the equilibrium growth rate using bisection. In each step I assume a growth rate, calculate the implied innovation rate and firms' discount rate. After solving the model I then check on the model implied growth rate and iterate until initial guess and model implied growth rate converge.

Simulation. Once I solved the model, I simulate data for a single firm with a number of R&D lines, which coincide in R&D productivity and, thus, in R&D employment. I set the number of R&D lines N such that the simulated data matches the average number of patents in my sample $N_P = 520$:

$$N = \frac{N_P}{z}.$$

In each period, I first determine the firms' optimal R&D policy using the policy function together with the associated R&D success probability. I then draw the number of successful inventions from a Bernoulli distribution using with the R&D success probability and number of R&D lines as parameters. For each of the successful inventions I then draw a step-size from the calibrated geometric distribution and record the implied patent valuation. Finally, I use the Markov process for the R&D productivity process to draw next periods R&D productivity.

I repeat this procedure until I have 100050 periods and discard the first 50 as burn-in. With the remaining data I follow the same steps as in my empirical exercise to calculate the relevant statistics.

C.5 Average and Aggregate Returns

The wage specification naturally implies that average are smaller than aggregate returns due to the positive correlation between R&D expenditure and returns. Table C.1 reports the associated numbers for the model and the data. In the data, the log aggregate return is about 20% larger than the average log return, while this gap is 15% in the model. Furthermore, the model implied returns are somewhat smaller than the returns on the data. For comparison, the model without frictions has aggregate and average log returns of $\log(1 + \phi) \approx 0.69$, which is far below the data equivalent. The model is thus able to speak to core facts about aggregate R&D performance.

Table C.1: Average and Aggregate Return in Data and Model

	Data	Model
Average ln Return on R&D	1.68	1.47
ln Aggregate Return on R&D	2.03	1.69

Notes: In the data, I calculate aggregate and average return first at the annual level before taking a simple average across years in the sample. In the model, I simulate a dataset with 100,000 observations and calculated both measures for the entire sample.

C.6 Perfect Price Discrimination

In this section I consider price discrimination across workers as an alternative perspective on R&D return dispersion. For this purpose, I ignore adjustment cost and work with the resulting static R&D choice model. As shown below, perfect price discrimination can result in R&D return dispersion, however, due to a very different mechanism. Under perfect price discrimination, firms equalize marginal benefit to unconstrained marginal cost, and, thus, marginal R&D returns are equalized as well, however, average and marginal return are no longer proportional. Thus, while average R&D returns can be still dispersed, it does not imply that resources are misallocated.

Lemma C.4. *Let $F(\ell) = \gamma \ell^{\frac{1}{1+\phi}}$ and assume that the cost function is given by*

$$C(\ell) = W \left(\int_0^{\ell} \left((1+\xi) \left(\frac{l}{L} \right)^{\xi} + \bar{\ell} \right) dl \right). \tag{C.13}$$

Then, firms take wages as given and set marginal benefit equal to marginal cost. Nonetheless, the Return on R&D is dispersed and given by

$$\frac{F(\ell^*)V}{C(\ell^*)} = \frac{1}{\varepsilon_F} \times \left(1 + \xi \times \frac{\left(\frac{\ell^*}{L}\right)^{\xi}}{\left(\frac{\ell^*}{L}\right)^{\xi} + \bar{\ell}} \right). \tag{C.14}$$

Proof. Integrating the cost function, we have

$$C(\ell) = \ell \times W \times \left(\left(\frac{\ell}{L} \right)^{\xi} + \bar{\ell} \right),$$

which is exactly the same formulation as in the main text. Resultingly, first order conditions coincide as does the specification of R&D returns. □

The difference in the models is not only of rhetorical importance, but also determines whether there is a market inefficiency assuming that wages are shaped by preferences. Under perfect price discrimination, firms equalize marginal cost and benefit such that the resulting allocation is efficiency. In contrast, under monopsony, firms equalize marginal benefits to marginal cost adjusted for a markdown reflecting their market power. The resulting allocation is inefficiency and a policy maker would want to subsidize firms with market power until marginal benefit and cost are equalized.

Equivalence of R&D returns and inventor supply elasticity under both models thus raises the larger question of how we can differentiate them in the data. The key difference is the price impact on infra-marginal inventors. In a world with perfect price discrimination, the infra-marginal wage is unaffected by labor demand. In contrast, under monopsony, infra-marginal wages are affected by decision on the margin. While my data is not able to directly shed light on this issue, the evidence in Seegmiller (2021) suggests that monopsony power is important. Using worker-level data, he documents wage change for incumbent worker of comparable magnitude to new hires in response to labor demand shocks. This evidence is at odds with a perfect price discrimination view if we consider incumbent workers to be infra-marginal and new hires to be marginal workers.

D A Result on Return Dispersion and Frictions

In this Appendix, I highlight one approach to quantifying the potential importance of R&D return dispersion in a simple growth model. The results are similar in spirit to Hsieh and Klenow (2009) and are further explored in the companion paper Lehr (2022). The main disadvantage of this approach is that it interprets all variation in R&D returns as frictions.

Theory. A unit mass of firms innovates with probability z_{it} each period, depending on their R&D efficiency φ_{it} and inventors hired ℓ_{it} via a decreasing returns to scale production function with scale elasticity $\frac{1}{1+\phi}$:

$$z_{it} = \varphi_{it} \ell_{it}^{\frac{1}{1+\phi}}.$$

Firms value innovation at expected value $\mathcal{V}_{it}$ and face common wage W_t. Input choices are distorted by exogenous wedge Δ_{it} such that their optimal inventor employment solves:

$$\frac{\partial z_{it}}{\partial \ell_{it}} \mathcal{V}_{it} = (1 + \Delta_{it}) \times W_t.$$

The wage W_t is determined via labor market clearing with mass of R&D workers $\mathcal{L}$:

$$\mathcal{L} = \int_0^1 \ell_{it} di.$$

Finally, the economic growth rate depends on innovation rates z_{it} and the growth impact of innovations $\lambda_{it} - 1$:

$$g_t = \int_0^1 z_{it}(\lambda_{it} - 1) di.$$

Proposition D.1. *Let the ratio of productivity impact to valuation be constant across firms, i.e. $V_{it} \propto \lambda_{it}$, and define a firm's R&D productivity as $\gamma_{it} \equiv \varphi_{it} \mathcal{V}_{it}$. We can express the economic growth-rate as the product of two factors:*

$$g_t = \tilde{g}_t \times \Xi_t. \tag{D.1}$$

The term $\tilde{g}_t$ captures the growth rate under the growth maximizing R&D worker allocation and is given by

$$\tilde{g}_t = \mathcal{L}^{\frac{1}{1+\phi}} \times \left(\int_0^1 \gamma_{it}^{\frac{1+\phi}{\phi}} di \right)^{\frac{\phi}{1+\phi}}. \tag{D.2}$$

The term Ξ_t captures the growth cost induced by frictions and can be interpreted as the *fraction of potential growth that is truly realized:*

$$\Xi_t = \frac{\int_0^1 \omega_{it}(1+\Delta_{it})^{-\frac{1}{\phi}} di}{\left(\int_0^1 \omega_{it}(1+\Delta_{it})^{-\frac{1+\phi}{\phi}} di\right)^{\frac{1}{1+\phi}}} \quad \textit{with} \quad \omega_{it} \equiv \frac{\gamma_{it}^{\frac{1+\phi}{\phi}}}{\int_0^1 \gamma_{it}^{\frac{1+\phi}{\phi}} di}. \tag{D.3}$$

Note that $\Xi_t \in (0,1]$ *and* $\Xi_t = 1$ *if* $\Delta_{it} = \Delta_t$.

Proof. The formulas follow by rearranging terms and solving for the growth rate. $\Xi_t \in (0,1]$ follows from Jensen's inequality since the denominator is a concave transformation of the nominator. □

Measurement. The proposition allows us to quantify the impact of R&D return dispersion within a basic endogenous growth framework. To estimate Ξ_t, we need three ingredients. Firstly, we need to fix ϕ and, as in the main text, I will set $\phi = 1$. Secondly, we need to measure Δ_{it}, which we can read off the return on R&D:

$$\text{R\&D Return}_{it} \equiv \frac{z_{it}\mathcal{V}_{it}}{W_t \ell_{it}} = (1+\phi) \times \Delta_{it}. \tag{D.4}$$

Note that the factor $1+\phi$ does not affect the calculations as the formula for Ξ_t is homogeneous of degree 0 in the scale of $1+\Delta_{it}$ and γ_{it}. This result is due to assuming a constant mass of R&D workers, such that heterogeneity in return and productivity only affects the allocation of R&D resources across firms. Finally, we need to measure R&D productivity. Rearranging firm order conditions, one can show that

$$\gamma_{it} \propto (1+\Delta_{it}) \times (W_t \ell_{it})^{\frac{\phi}{1+\phi}}, \tag{D.5}$$

which is sufficient to pin down relative productivity, and, thus, sufficient to construct Ξ_t.

Results. We thus have all the requirement ingredients and can implement the formulas. Column(1) in Table D.1 reports the results. Without adjustments, the model estimates an aggregate R&D allocation efficiency around 60%. Taken at face value, the estimate implies that the growth rate would be $1/0.6 - 1 = 67\%$ larger. Once we make the measurement adjustment lined out in Section 3, our estimate increase to 70% efficiency, implying a potential gain from reducing Return on R&D dispersion around 40%. Against a baseline growth rate of 1.5%, these estimates suggest a potential gain of 1 and 0.6 p.p. annual growth.

Table D.1: Allocative Efficiency Estimates

Return on R&D	Main	Measurement Error 15%	Measurement Error 30%
Baseline	60.2%	69.2%	78.1%
Adjusted	71.7%	78.5%	84.8%

Note: Estimates following Proposition D.1 assuming $\phi = 1$. Measurement error adjustments shrink log R&D by 1 minus adjustment fraction.

For comparison, Hsieh and Klenow (2009) estimate that US productivity, and thus production, could be 40% larger without dispersion in the total revenue productivity, which is conceptually similar to R&D returns. Similarly, Berger et al. (2022) estimate that US output could be 21% larger in absence of monopsony power. My estimates are on the same order of magnitude, however, they concern the growth rate and not productivity level. This difference has important welfare implications as welfare tends to be more sensitive to productivity growth rather than level due to its cumulative nature.

Discussion. One caveat with this approach is that we have to interpret all variation in the R&D returns as being driven by Δ_{it}. For example, measurement error raises R&D return dispersion, which in turn mechanically leads to lower estimates of Ξ_t. I highlight this challenge in column (2) and (3) in Table D.1, where I assume that measurement error constitutes 15% and 30% of the variation respectively. Mechanically, this assumption pushes up the estimated R&D allocation efficiency. Another approach to dealing with measurement error is to focus on changes over time. I explore this in detail in Lehr (2022) and find that indeed the dispersion in R&D returns has risen since 1975. Through the lens of the model, this suggest that misallocation has worsened, potentially contributing to the growth slowdown documented in Syverson (2017).

E Top 50 and Bottom 50 Firms by Return on R&D

Table E.1: Top and Bottom Companies by average Return on R&D

Rank	Company Name	Avg. ln Return on R&D
1	BJ SERVICES CO	3.88
2	INTUITIVE SURGICAL INC	3.76
3	AT&T INC	3.68
4	CAMERON INTERNATIONAL CORP	3.55
5	ILLINOIS TOOL WORKS	3.45
6	SALESFORCE.COM INC	3.42
7	WEATHERFORD INTL PLC	3.38
8	CREE INC	3.33
9	ARCHER-DANIELS-MIDLAND CO	3.33
10	INTL PAPER CO	3.32
11	MOBIL CORP	3.22
12	HALLIBURTON CO	3.22
13	UNOCAL CORP	3.20
14	DELL TECHNOLOGIES INC	3.19
15	CONOCOPHILLIPS	3.18
16	EXXON MOBIL CORP	3.16
17	ALIGN TECHNOLOGY INC	3.16
18	DEXCOM INC	3.14
19	BAKER HUGHES INC	3.14
20	QUALCOMM INC	3.13
21	OCCIDENTAL PETROLEUM CORP	2.99
22	ALZA CORP	2.98
23	TEXACO INC	2.96
24	ATLANTIC RICHFIELD CO	2.95
25	CHEVRON CORP	2.95

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. See Section 3 for details on the data construction.

Table E.2: Top and Bottom Companies by average Return on R&D (continued)

Rank	Company Name	Avg. ln Return on R&D
26	BLACKBERRY LTD	2.95
27	AMOCO CORP	2.95
28	LINDSAY CORP	2.95
29	RED HAT INC	2.94
30	U S SURGICAL CORP	2.94
31	RESMED INC	2.91
32	AKAMAI TECHNOLOGIES INC	2.89
33	STANDARD OIL CO	2.89
34	ALTERA CORP	2.88
35	MICRON TECHNOLOGY INC	2.87
36	UNIVERSAL DISPLAY CORP	2.87
37	SUNPOWER CORP	2.85
38	FORTINET INC	2.84
39	SYMBOL TECHNOLOGIES	2.83
40	BEAM INC	2.83
41	ECOLAB INC	2.81
42	BROADCOM INC	2.81
43	COOPER INDUSTRIES PLC	2.79
44	ALPHABET INC	2.79
45	SANDISK CORP	2.77
46	WEST PHARMACEUTICAL SVSC INC	2.74
47	APPLE INC	2.74
48	ACUITY BRANDS INC	2.73
49	DIGIMARC CORP	2.73
50	KERR-MCGEE CORP	2.72

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. See Section 3 for details on the data construction.

Table E.3: Top and Bottom Companies by average Return on R&D (continued)

Rank	Company Name	Avg. ln Return on R&D
419	AEROQUIP-VICKERS INC	0.67
420	AEROJET ROCKETDYNE HOLDINGS	0.67
421	SILICON GRAPHICS INC	0.66
422	AMERICAN AXLE & MFG HOLDINGS	0.65
423	COHERENT INC	0.65
424	AVID TECHNOLOGY INC	0.65
425	ITRON INC	0.65
426	TELLABS INC	0.64
427	GOULD INC	0.63
428	MILACRON INC	0.60
429	RIGEL PHARMACEUTICALS INC	0.57
430	BECKMAN COULTER INC	0.55
431	MAXYGEN INC	0.55
432	HASBRO INC	0.54
433	MICROVISION INC	0.53
434	FIRESTONE TIRE & RUBBER CO	0.52
435	APPLIED MICRO CIRCUITS CORP	0.51
436	MODINE MANUFACTURING CO	0.48
437	FORD MOTOR CO	0.46
438	DIGITAL EQUIPMENT	0.46
439	CELANESE CORP-OLD	0.46
440	SCOTT TECHNOLOGIES INC	0.45
441	AXCELIS TECHNOLOGIES INC	0.42
442	ANALOGIC CORP	0.41
443	QUANTUM CORP	0.40

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. See Section 3 for details on the data construction.

Table E.4: Top and Bottom Companies by average Return on R&D (continued)

Rank	Company Name	Avg. ln Return on R&D
444	AMDOCS	0.39
445	DATA GENERAL CORP	0.39
446	SPERRY CORP	0.37
447	ELECTRO SCIENTIFIC INDS INC	0.35
448	NAVISTAR INTERNATIONAL CORP	0.31
449	MAXTOR CORP	0.30
450	QLOGIC CORP	0.29
451	MCDONNELL DOUGLAS CORP	0.29
452	TANDEM COMPUTERS INC	0.29
453	TELECOMMUNICATION SYS INC	0.24
454	ROBINS (A.H.) CO	0.23
455	SPANSION INC	0.17
456	ELECTRONICS FOR IMAGING INC	0.15
457	EXTREME NETWORKS INC	0.13
458	WANG LABS INC	0.09
459	BIO-RAD LABORATORIES INC	0.08
460	DAY INTERNATIONAL INC	0.08
461	ROGERS CORP	-0.08
462	SMITH (A.O.)	-0.09
463	GENERAL MOTORS CO	-0.14
464	AMDAHL CORP	-0.30
465	VISTEON CORP	-0.34
466	MENTOR GRAPHICS CORP	-0.35
467	DE SOTO INC	-0.37
468	DONNELLY CORP	-0.40

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. See Section 3 for details on the data construction.

Table E.5: Top and Bottom Companies by average adjusted Return on R&D

Rank	Company Name	Avg. ln Return on R&D
1	INTUITIVE SURGICAL INC	3.82
2	DEXCOM INC	3.64
3	DIGIMARC CORP	3.58
4	ILLINOIS TOOL WORKS	3.53
5	CREE INC	3.52
6	BROADCOM INC	3.51
7	SALESFORCE.COM INC	3.47
8	FORTINET INC	3.45
9	AT&T INC	3.34
10	QUALCOMM INC	3.34
11	GENTEX CORP	3.20
12	MICRON TECHNOLOGY INC	3.17
13	ECOLAB INC	3.17
14	F5 NETWORKS INC	3.16
15	SUNPOWER CORP	3.14
16	UNIVERSAL DISPLAY CORP	3.14
17	RED HAT INC	3.12
18	RESMED INC	3.12
19	ILLUMINA INC	3.11
20	CAMERON INTERNATIONAL CORP	3.06
21	MICROSOFT CORP	3.02
22	BLACKBERRY LTD	2.96
23	ALTERA CORP	2.95
24	AIR PRODUCTS & CHEMICALS INC	2.94
25	DELL TECHNOLOGIES INC	2.91

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. Adjustments include (1) winsorizing patent valuations, (2) knowledge capital, (3) NAICS3$\times$ Year effects, (4) amenities, and (5) acquisitions. See Section 3 and B.1 for details.

Table E.6: Top and Bottom Companies by average adjusted Return on R&D (continued)

Rank	Company Name	Avg. ln Return on R&D
26	LINDSAY CORP	2.91
27	MASIMO CORP	2.88
28	ALZA CORP	2.88
29	APPLE INC	2.86
30	SANDISK CORP	2.86
31	VIASAT INC	2.82
32	FUELCELL ENERGY INC	2.81
33	ALIGN TECHNOLOGY INC	2.81
34	NVIDIA CORP	2.80
35	VMWARE INC -CL A	2.80
36	XILINX INC	2.79
37	SYMBOL TECHNOLOGIES	2.79
38	ALCOA INC	2.78
39	WATERS CORP	2.74
40	NETLOGIC MICROSYSTEMS INC	2.70
41	LIFE TECHNOLOGIES CORP	2.69
42	PITNEY BOWES INC	2.65
43	CORNING INC	2.60
44	U S SURGICAL CORP	2.59
45	AMKOR TECHNOLOGY INC	2.56
46	VERTEX PHARMACEUTICALS INC	2.54
47	LINEAR TECHNOLOGY CORP	2.54
48	PROCTER & GAMBLE CO	2.54
49	AKAMAI TECHNOLOGIES INC	2.53
50	COLGATE-PALMOLIVE CO	2.52

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. Adjustments include (1) winsorizing patent valuations, (2) knowledge capital, (3) NAICS3× Year effects, (4) amenities, and (5) acquisitions. See Section 3 and B.1 for details.

Table E.7: Top and Bottom Companies by average adjusted Return on R&D (continued)

Rank	Company Name	Avg. ln Return on R&D
419	AGCO CORP	0.88
420	TORO CO	0.86
421	CORDIS CORP	0.86
422	FORD MOTOR CO	0.83
423	TANDEM COMPUTERS INC	0.82
424	MILACRON INC	0.81
425	VEECO INSTRUMENTS INC	0.81
426	DENNISON MFG CO	0.78
427	MERITOR INC	0.77
428	MAXYGEN INC	0.76
429	INTERMEC INC	0.76
430	RIGEL PHARMACEUTICALS INC	0.75
431	SURGALIGN HOLDINGS INC	0.73
432	ACTEL CORP	0.72
433	ELECTRO SCIENTIFIC INDS INC	0.71
434	APPLIED MICRO CIRCUITS CORP	0.70
435	G-I HOLDINGS INC	0.70
436	CA INC	0.69
437	MODINE MANUFACTURING CO	0.69
438	COHERENT INC	0.68
439	MICROSTRATEGY INC	0.67
440	QLOGIC CORP	0.67
441	TELLABS INC	0.62
442	ZENITH ELECTRONICS CORP	0.60
443	MENTOR GRAPHICS CORP	0.60

Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. Adjustments include (1) winsorizing patent valuations, (2) knowledge capital, (3) NAICS3× Year effects, (4) amenities, and (5) acquisitions. See Section 3 and B.1 for details.

Table E.8: Top and Bottom Companies by average adjusted Return on R&D (continued)

Rank	Company Name	Avg. ln Return on R&D
444	QUANTUM CORP	0.55
445	AVID TECHNOLOGY INC	0.55
446	CELANESE CORP-OLD	0.53
447	NAVISTAR INTERNATIONAL CORP	0.50
448	BECKMAN COULTER INC	0.50
449	LUBRIZOL CORP	0.49
450	CONEXANT SYSTEMS INC	0.46
451	AMDAHL CORP	0.41
452	SPANSION INC	0.41
453	ROGERS CORP	0.39
454	EXTREME NETWORKS INC	0.36
455	ANALOGIC CORP	0.35
456	MAXTOR CORP	0.33
457	DONNELLY CORP	0.33
458	GENERAL MOTORS CO	0.24
459	AXCELIS TECHNOLOGIES INC	0.21
460	ROBINS (A.H.) CO	0.20
461	ELECTRONICS FOR IMAGING INC	0.20
462	STEEL EXCEL INC	0.17
463	HASBRO INC	-0.03
464	BIO-RAD LABORATORIES INC	-0.06
465	AT&T CORP	-0.09
466	DE SOTO INC	-0.20
467	VISTEON CORP	-0.23
468	SMITH (A.O.)	-0.27

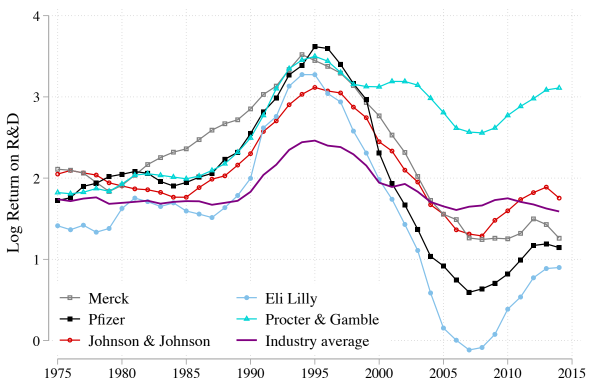
Note: This table reports the company names of firms with the best and worst average ln Return on R&D in the sample. I restrict the list to firms with at least 10 observations. Adjustments include (1) winsorizing patent valuations, (2) knowledge capital, (3) NAICS3$\times$ Year effects, (4) amenities, and (5) acquisitions. See Section 3 and B.1 for details.

F Case Studies

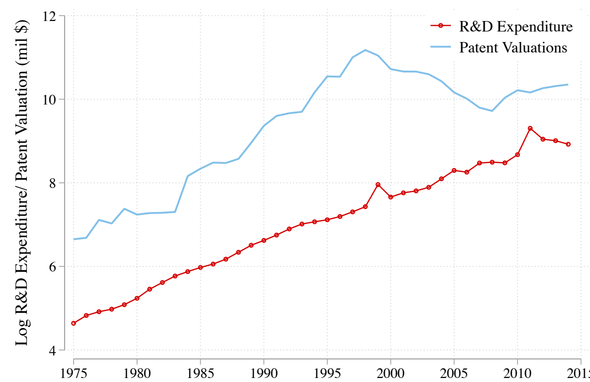
In this section I highlight selected case studies that provide further context on the dynamics and origin of R&D return dispersion. The first case study focuses on Merck and the Pharma industry, shedding a light on some of the most innovative firms in the economy in an industry where patent rights are key to guarding innovation from competitors. The second case study takes a look at the natural resource industry, which, perhaps surprisingly, has earned larger R&D returns than any other industry.

F.1 Merck & the Pharma Industry

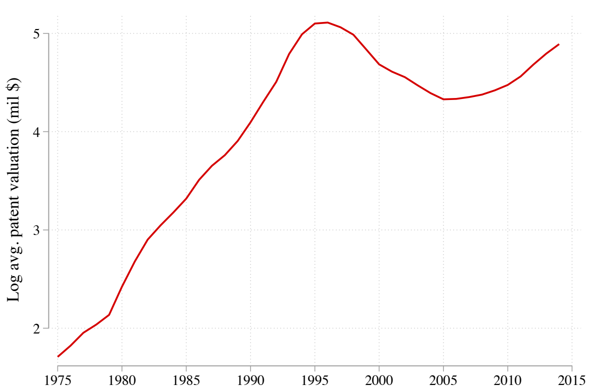
Figure F.1: Merck's R&D Performance 1975-2014



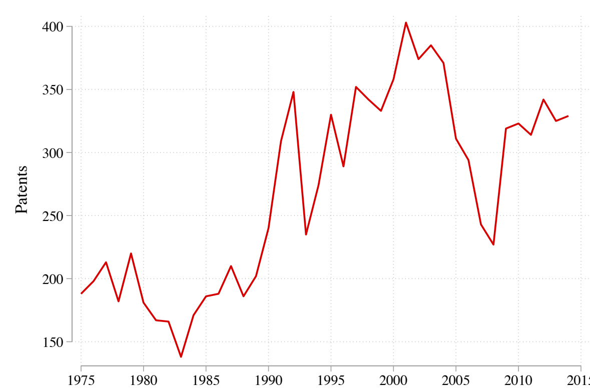
(a) Return on R&D



(b) R&D Expenditure & Patent Valuations



(c) Average Patent Valuation



(d) Patents

Notes: Panel (a) plots the 5-year return on R&D for selected firms within chemical manufacturing (NAICS 325) as well as an industry average for firms with at least 20 active years within the sample. Panel (b)-(d) focus on Merck & Co only. Panel (b) plots annual R&D expenditure and patent valuations. Panel (c) plots the average patent valuation at the 5-year level. Panel (d) plots the annual Cnumber of patents.

Panel (a) in Figure F.1 plots the evolution of R&D returns for Merck, important competitors, and the industry average. R&D returns are relatively constant until the late 1980s, where they begin to rise. Returns peak around 1995 and, for all but Proctor & Gamble, subsequently return to their previous level or lower. R&D returns are notably more dispersed post 2005 than in previous decades.

Panels (b)-(d) take a closer look at Merck in particular. In Panel (b) I plot the two components of R&D returns, patent valuations and R&D expenditure, separately. The emerging pattern is one of an essentially constant growth rate of R&D expenditure over the entire sample, while patent valuation drive fluctuations in R&D returns by first accelerating in the early parts of the same and subsequently declining below their initial trend. Panels (c) and (d) reveals that the evolution of patent valuation is driven both by rising patent valuations as well as rising patent counts. Annual patenting is centered around 175 for the 1975 to 1990 period before jumping to a new average level around 300. Patent valuation grow smoothly from 1975 to 1995 and subsequently stabilize.

The emerging patterns suggest that the evolution of R&D returns for Merck is driven partly by a failure to respond to rising innovation output by increasing R&D expenditure and vice versa. Given the year-to-year stability in returns, it appears unlikely that this is driven by perceived uncertainty around the value of innovation. The stability of R&D expenditure across years further raises the question as to the underlying decision making process and, potentially, highlights the importance of adjustment cost, e.g. due to the scarcity of talent, in the R&D process.

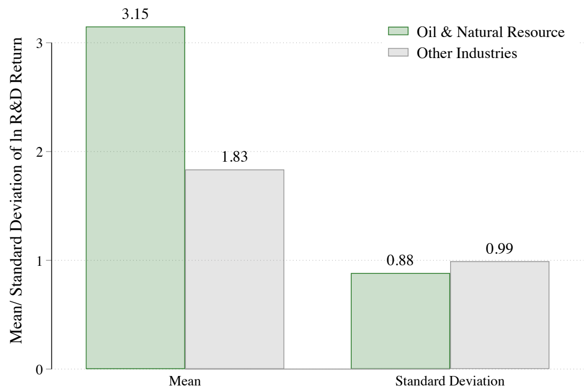
F.2 Exxon and the Natural Resource Industry

Another interesting case is the natural resource industry. As show in Panel (a) in Figure F.2, the average firm in the industry earns a significantly higher return than firms in other industries, however, the dispersion within the industry is slightly lower than outside the industry.

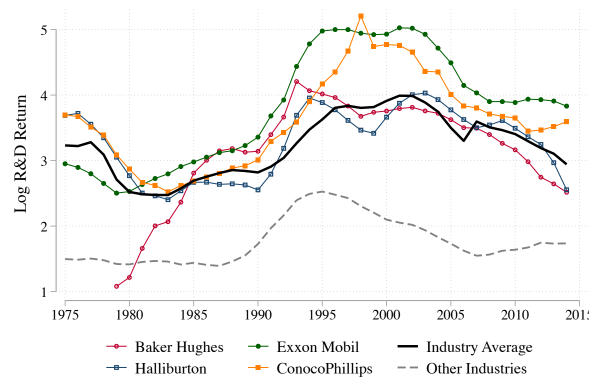
Panel (b) plots the evolution of R&D returns for selected firms in the industry, the industry average, and the average of firms outside the industry. Returns are initially stable until 1990 and subsequently peak around 2000 before returning to pre-peak levels around 2005. Interestingly, the ranking of R&D returns across the four competitors shown is very stable across years with Exxon Mobil earning the highest returns in most years.

Panels (c) and (d) take a closer look at Exxon. Panel (d) plots the evolution of R&D expenditure and patent valuations. The figure reveals very stable R&D expenditure and patent valuation however at different growth rates. Furthermore, patent valuations experience a temporary peak around the 2000s, however, the peak is much stronger for Exxon Mobil and ConocoPhillips than for their competitors.

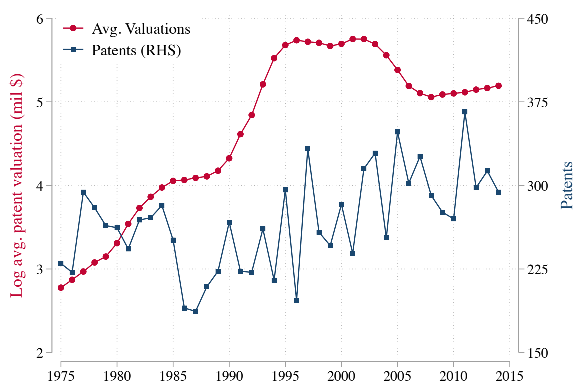
Figure F.2: Exxon's R&D Performance 1975-2014



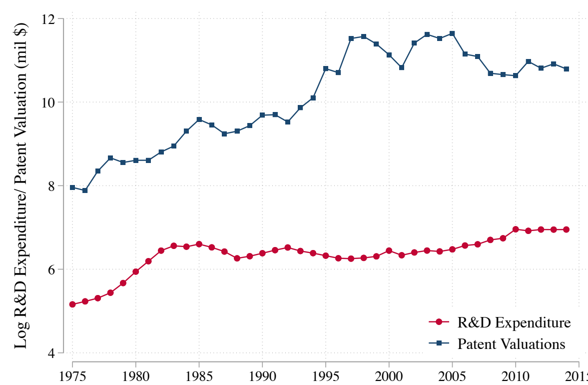
(a) Summary Statistics



(b) R&D Returns



(c) Patents & Their Valuation



(d) R&D Expenditure & Patent Valuations

Notes: Panel (a) plots the average return and standard deviation thereof within the industry and outside of the industry. Panel (b) plots the 5-year return on R&D for selected firms within natural resource industry (NAICS 211,213, and 324) as well as an industry average for firms with at least 20 active years within the sample. Panel (c)-(d) focus on Exxon Mobile only. Panel (d) plots annual R&D expenditure and patent valuations. Panel (c) plots the average patent valuation at the 5-year level and annual patents. Panel (d) plots the annual number of patents.