The Relationship Between Productivity and R&D : Uncertainty and Heterogeneity*

Chapter 2: Version 1.0

Kevin Andrew†

September 9, 2019

Abstract

In this paper I examine the effects of R&D spending on productivity over the life cycle of a firm. I do this by estimating a TFP process which depends on age, lagged productivity and R&D spending. My principal contribution is to account for the amount of uncertainty inherent in R&D in an explicit way. I do this by estimating a skedastic function. I find that firms which do more R&D have more volatile productivity and that this effect is heterogeneous over the firm life cycle. In particular, the uncertainty of R&D is declining as a firm ages. As well, firms which have higher productivity have more uncertainty research outcomes.

Journal of Economic Literature Classification:

Keywords: Firm Dynamics, Productivity, Research and Development, Production Function Estimation, Knowledge Capital

---

* This research was support by the Social Sciences and Humanities Research Council of Canada and the John Deutsch Institute Doctoral Stipend. All errors are my own.

† Queen's University, Department of Economics, Kingston, Ontario, K7L 3N6.
Email: andrewk@econ.queensu.ca

1 Introduction

It is well established that a nation's level of aggregate productivity is of first order importance for living standards. Aggregate productivity encapsulates millions of individual decisions of producers. These include choices about whether or not to shut down the business and how much effort and money to invest in research and development. Firm level data has shown that movements in aggregate productivity often mask micro-level churning which is an order of magnitude larger. The relationship between R&D and productivity has been studied extensively within economics in the last 50 years.[1] My paper contributes to this literature by addressing the impact of R&D on productivity, specifically focusing on uncertainty and heterogeneity in the research process.

Uncertainty. Conducting R&D is inherently different than investing in tangible capital. In particular, R&D investment is highly uncertain. The risks inherent in investing in tangible capital include cost overruns, delays and disruption to ongoing business. While these risks are non-negligable, they are dwarfed by the uncertainty inherent in R&D . Many investments in R&D are abandoned altogether. Furthermore, there are many cases where firms are unable to implement good ideas. In this paper I find that the more R&D a firm does, the more uncertain productivity growth is.

Heterogeneity. There is reason to believe that the nature of R&D projects that a firm undertakes may be heterogeneous with respect to age as well as already attained productivity. For instance, an established company may choose to do R&D projects which are inherently different from new entrants. I find that the uncertainty of R&D projects is decreasing over the life cycle. That is, the uncertainty inherent in the R&D process decreases as firms age.[2]

The strategy I use to isolate these interesting interactions is to estimate a production function using data on sales, employment, capital, materials and R&D spending. I take my data from the Compustat database. As part of my empirical strategy I estimate a TFP process which depends on R&D expenditures as well as age and lagged productivity. I also estimate a function which documents the conditional heteroskedasticity of TFP volatility. In this way, R&D can influence both the mean and variance of next periods productivity. It turns out that the variance effects are most significant.

[1] The empirical study of production functions with knowledge capital was pioneered by Griliches et al. (1979). Doraszelski and Jaumandreu (2013) represent the state of the art and my paper builds off their work. The endogenous growth literature stemming from Romer (1990) and Aghion and Howitt (1992) emphasize the role of R&D as a mechanism for the aggregate growth the economy. Acemoglu et al. (2018) is a recent paper in this literature which employs firm level data. The empirical literature on productivity dynamics is summarized in Bartelsman and Doms (2000), Foster et al. (2001) and Syverson (2011).

[2] **I am having trouble writing these results intuitively at this point, so to sum up the last two paragraphs in a specific manner...** If the volatility of the TFP shock is a function of the amount of R&D and age, the cross derivatives satisfy:

$$\frac{\partial^2 \sigma(r,a)}{\partial r \partial a} < 0 \qquad \frac{\partial \sigma(r,a)}{\partial r} > 0 \qquad \frac{\partial \sigma(r,a)}{\partial a} < 0$$

Methodological Contribution 1 In this paper I make two main methodological contributions. First, I extend the methodology of Doraszelski and Jaumandreu (2013) to a more general setting that closer matches firm level data available in representative samples. In particular, Doraszelski and Jaumandreu (2013) rely on firm level variation in prices and wages for identification. My method relies instead on inverting the static materials demand function as in Levinsohn and Petrin (2003). In the context of the literature on production function estimation, I combine these two approaches to estimate my production function and controlled Markov process without wage and price data. I also control for endogenous selection. To my knowledge, this strategy has not been used before in the literature.[3]

Methodological Contribution 2 My second methodological contribution is to include conditional heteroskedacitity in R&D . I do this by estimating a skedastic function, which can depend on age, productivity and research. This allows me to study the effects of R&D spending on the volatility of TFP. Young firms are known to have more volatile employment growth rates than mature firms. I want to systematically study how research and productivity growth might be related to this finding. In order to do that I need to have an empirical strategy for identifying the effects of age and research on the variance of productivity growth. This skedastic function allows me to do that. I show how including conditional heteroskedasticity does not affect the population moment conditions used in the literature.

Literature 1: Knowledge Capital The importance of knowledge capital in production of goods and services has been studied extensively, following Griliches et al. (1979). Hall et al. (2010) surveys this literature in the context of measuring the returns to R&D . There are measurement issues inherent in constructing the stock of knowledge capital. Most prominent is the choice of depreciation rate of knowledge.[4] Doraszelski and Jaumandreu (2013) get around this problem by focusing on the flows of R&D spending at the firm level. I follow this approach in this paper. It is generally assumed that knowledge capital is accumulated similar to tangible capital. I find that the accumulation of knowledge is highly uncertain which suggests that the law of motion for intangible capital should include some randomness. In this sense my work is complementary but not identical to the knowledge capital literature.

Literature 2: The Firm Life Cycle With the increasing availability of high quality firm level data, many recent studies have documented salient facts about the life cycle of firms.[5] Collectively these are referred to as the "up or out" dynamic. In sum, young firms grow faster on average, have more uncertain growth rates and are more likely to

[3]The paper of Gandhi et al. (2018) calls into question the practice of estimating a value added production function as I do in this paper. I believe I can combine my approach with theirs to develop a gross production function estimator. For now, I just assume that the elasticity of substitution between materials and other inputs is 0.

[4]This parameter is not well identified from panel data and it is not clear that it should even be constant over time. If the creative destruction rate varies over time then the depreciation of knowledge is not constant.

[5]Much of this research is summarized in Decker et al. (2014) in the context of the decline in business dynamism.

exit. Furthermore, it is found that this higher uncertainty is not symmetric. That is, the growth rate distribution of young firms is right skewed. My empirical methodology allows me to examine these facts through the lens of my model of productivity. I find that TFP growth does not exhibit a strong trend over the life cycle, but the variance of TFP growth is much higher for young firms. I find that doing more R&D increases the dispersion of productivity growth and that the uncertainty of R&D decreases over the life cycle. All of these findings add context to the interpretation of the "up or out" dynamics of young firms. **I tentatively conclude that the decline in variance in firm growth rates over the life cycle is related to both my R&D mechanism and an exogenous age effect. The declining mean growth rates appear to be related to convex adjustment costs to tangible capital accumulation. The interaction of non-convex capital adjustment costs, endogenous selection and the choice of R&D projects may explain this "up or out" dynamic for the firm life cycle.**

Literature 3: Endogenous Growth Lastly, this work is related to recent work in the endogenous growth literature which uses firm level data. In particular, Akcigit and Kerr (2018) suggest that the composition of R&D is different across the firm size distribution. They argue that larger firms conduct more internal innovation, which is aimed at improving existing product lines. Small firms conduct more external innovation, aimed at discovering new product lines. Entrants perform only external innovation. I do not explicitly model the type of R&D , but I do allow interactions between productivity, age and R&D . I find that these interactions are important for the volatility of TFP, but not exactly in the way that the Akcigit and Kerr (2018) mechanism would suggest. I find that as firms age, their R&D spending is less uncertain. I also find that firms with higher TFP have more uncertain R&D , inconsistent with the external vs. internal dichotomy. [6]

Thesis Plan 1 The finding that R&D is uncertain motivates the rest of my thesis. In Chapter 3 I will examine the theoretical implications of this uncertainty in the context of a simple model of selection and firm growth. In particular, I will show how endogenous R&D choices in the presence of irreversibility of capital investment interact with selection. This simple mechanism can account for the qualitative aspects of the up or out dynamic of the firm life cycle. In particular, if a firm knows it will make an optimal exit decision in the future, a firm with a lower capital stock will choose a more uncertain outcome. Given that R&D increases uncertainty, this mechanism can explain why young firms who have yet to accumulate much tangible capital have higher average growth rates, higher uncertainty of those growth rates, higher skewness of growth rates and higher exit rates.

Thesis Plan 2 In Chapter 4 I will place the mechanism outlined in the previous paragraph in a general equilibrium model of firm dynamics with endogenous R&D choices. I will quantify the importance of this mechanism for productivity growth over the life cycle and consider the effects of an exogenous decline in the entry rate due to a fall in the supply of potential entrepreneurs. In this model there is not market failure, so I will

[6] **Need to think more about the implications of this.**

essentially be viewing the mechanics of productivity growth in the context of changing demographics. I will compare my findings with those of Karahan et al. (2019) who introduced this potential cause of the decline in entry.

Thesis Plan 3 Chapter 5 will introduce knowledge spillovers, and thus a role for innovation policy. This is an important step as this mechanism is often emphasized in the endogenous growth literature. Furthermore, Akcigit and Ates (2019) argues that changes in the structure of the patent system may be behind the decline in entry. In order to talk about innovation policy in a meaningful way I have to first identify the magnitude of spillovers. These estimates will discipline my quantitative exploration of the importance of changes in the patent system in recent decades for business dynamism and productivity.

Plan of the paper. The paper proceeds as follows. Section 2 introduces a model of R&D and productivity growth as well as my estimation procedure. I discuss the different specifications for TFP I consider in this papre. Section 3 introduces the dataset, data construction procedure and sample selection. I compare my non-representative sample with findings from the entire universe of US firms. Section 4 presents the main findings of the paper. Section 5 is unfinished and will identify knowledge spillovers. Section 6 is unfinished and will include robustness. I have performed some robustness at this point and in other cases I suggest ways I could check the stability of my main conclusions. Section 7 concludes.

2 Empirical Methodology Data

The purpose of this section is to introduce my empirical methodology and to compare it to the existing literature. This will make clear the exact assumptions that my strategy relies upon as well as the things which make this study unique. In Section 2.1 I start by introducing a model of R&D where productivity follows a controlled Markov process. I discuss the assumptions which are necessary for identifying this stochastic process as well as my strategy for implementing the estimator. In Section 2.2 I compare my estimator to the existing literature on knowledge capital and production function estimation. This section sheds light on the reasons for various modelling choices. Lastly, I discuss the exact models that I will estimate in Section 2.3.

2.1 The Model and Empirical Methodology

2.1.1 A Model of R&D and Productivity

The unit of analysis is an individual firm. A firm transforms inputs into outputs via its production function.

Definition 1. ***The Gross Production Function:***

$$Y_{i,j,t}^{G} = \min\{\beta_0 K_{i,j,t}^{\beta_k} L_{i,j,t}^{\beta_l} e^{\omega_{i,j,t}}, \beta_m M_{i,j,t}\} e^{\varepsilon_{i,j,t}}$$

Where:

$$\omega_{i,j,t} = \iota_j + \tau_t + z_{i,j,t}$$

Where i denotes a firm, j denotes an industry and t denotes a year. The empirical strategy differentiates between "dynamic" inputs, $\mathbf{x}_{i,j,t}$ and "static" inputs, $\mathbf{m}_{i,j,t}$. For the purposes of this study, the vector of dynamic inputs includes labour and capital, which must be chosen one period ahead. These are subject to adjustment costs. In this paper I estimate a value added production function. The reason why I assume a Leontief specification for gross output is the identification concerns raised by Ackerberg et al. (2015) and Gandhi et al. (2018). In particular, variable input elasticities are not identified in a value added production function. This is why I assume that labour and capital are dynamic inputs and proceed with the Leontief specification for gross output. I can then define the value added production function as gross output minus materials input. I will invert a materials demand function following the methodology of Levinsohn and Petrin (2003).[7]

I have included industry and year dummies. My approach is different from the industrial organization literature, which often seeks to identify different production function elasticities for different industries. My focus is on identifying a single firm level productivity process. What I will define as value added productivity for the purposes of this project is $z_{i,j,t}$. **Further robustness could be done on how estimates vary across industries.**

The productivity residual is split into two components: one which the firm observes (but I do not), $z_{i,j,t}$, and the other which neither observe, $\varepsilon_{i,j,t}$. Accounting for the endogeneity inherent in this assumption will be addressed. I also control for the endogenous exit decision of the firm. This turns out to be important for interpreting the stochastic process for TFP.

I follow Doraszelski and Jaumandreu (2013) and assume that the process for productivity is controlled by the decision making of the firm. In particular, choices about R&D spending influence the evolution of productivity over time. I can also allow factors such as aggregate R&D , firm age and other important characteristics influence the TFP process. I collapse all these variables in the vector $\mathbf{w}_{i,j,t}$ and define the controlled Markov Process as follows:

Definition 2. ***Controlled Markov Process:***

$$z_{i,j,t+1} = g(z_{i,j,t}, r_{i,j,t}, \mathbf{w}_{i,j,t}) + \sigma(r_{i,j,t}, \mathbf{w}_{i,j,t})\xi_{i,j,t+1}$$

Notice that I have allowed the volatility of the productivity process to be endogenous as well. Given the inherent uncertainty characteristic of the R&D process I think this is an important extension of the model. I believe I am the first to systematically estimate this type of process. In particular, Doraszelski and Jaumandreu (2013) emphasize the non-linearities in the R&D accumulation process. However, they do not focus on firm age or a skedastic function. I also extend their ideas into settings where there is no

[7] I have also estimated the model with firm fixed effects and time fixed effects. My results did not change very much. I cannot include firm, industry and time fixed effects due to collinearity.

variation in input prices. Given that most representative firm level datasets do not feature such variation, this makes the framework more broadly applicable.[8]

The additional model primitives are standard. Investment in dynamic inputs is subject to adjustment costs. Importantly, corner solutions for investment in these inputs are admissable due to the materials inversion. This is important as there is a large literature arguing that firm investment behaviour is lumpy. I also assume convex costs for R&D which include fixed costs. This leads to both an intensive and extensive margin for the R&D policy which will have implications for the interpretation of the results. Profits are defined as:

$$\pi(z, \mathbf{x}) = y - \mathbf{w}_x^T \mathbf{x}$$

Where $\mathbf{w}_x$ are a vector of input prices for the dynamic factors of production.

I have now described the model primitives which allow me to state the firms dynamic program and policy functions. I drop time and industry subscripts.

$$V_a(z, \mathbf{x}) = \max\{\Phi, [\max_{\mathbf{x} \in F(\mathbf{x}), r \geq 0} \quad \pi(z, \mathbf{x}) - \Psi_x(\mathbf{x}') - \Psi_r(r) + \beta \int V_{a+1}(z', \mathbf{x}') dG(z'|z, r, \mathbf{w})]\}$$

Where $\Psi_x(\mathbf{x}')$ represent adjustment costs for dynamic inputs and $\Psi_r(r)$ denote adjustment costs for R&D . The salvage value, Φ induces exit. The value function is non-decreasing in z which leads to an optimal-stopping rule for exit where firms exit only if productivity falls below a threshold, $\bar{z}(\mathbf{x})$.[9] Given that this problem is standard I don't discuss it further.[10]

The optimal R&D policy will be age dependent and will depend on productivity as well as the dynamic state vector, $\mathbf{x}$. I constrain R&D to be greater than 0 and allow for fixed costs. This implies that the R&D policy is of the form:

$$r_a(z, \mathbf{x}) = \begin{cases} \tilde{r}_a(z, \mathbf{x}) & (z, \mathbf{x}) \in \mathcal{X} \\ 0 & \end{cases}$$

Where $\mathcal{X}$ is the space where R&D is positive.

For the purposes of this paper it will not be critical to characterize all these policy functions sharply. In the next section I will provide the relevant assumptions which justify my empirical strategy. Then I will outline the multi-step procedure I use to identify the productivity process.

[8] I estimate my production function using Compustat Data, but it would be feasible to do using Longitudinal Business Database (LBD) data in the US and T2-LEAP data in Canada.

[9] I haven't actually proved this. In fact, the non-linearities may break the monotonicity result. I am in the process of solving the firms problem for Chapter 3 of my thesis and will know soon whether monotonicity holds at the estimated parameter values.

[10] Some notes: it is possible to allow adjustment costs to depend on current $\mathbf{x}$. The non-decreasing property of the value is only important for versions of the model where I control for exit. Otherwise, materials demand is sufficient.

2.1.2 Assumptions

I will begin by making two major assumptions, which I number. First, I require an invertible materials demand function as in Levinsohn and Petrin (2003). This allows me to control for the endogeneity of the input decisions in estimating the production function.

Assumption 1. ***Invertible Materials Demand:*** *I assume that materials is a function of the state and that it is strictly monotonic in productivity:*[11]

$$\boldsymbol{m}(z, \mathbf{x})$$

Under **Assumption 1** it is immediately the case that productivity is an unknown function of materials and the dynamic inputs. I call this function:

$$z = h(\mathbf{m}, \mathbf{x})$$

I am also explicit about the information structure. This helps in identifying potential instruments to be used in identifying the parameters.

Assumption 2. ***Information Structure and Timing:*** *At time t the firm has an information set $\mathcal{I}_t$ which includes current and past productivity, $\{z_{i,j,t}\}_0^t$. It does not include future productivity shocks. The transitory shocks and innovations to productivity satisfy:*

$$\mathbb{E}[\varepsilon_{i,j,t} + \sigma(z_{i,j,t}, r_{i,j,t}, \mathbf{w}_{i,j,t})\xi_{i,j,t+1}|\mathcal{I}_t] = 0$$

I also assume that the innovations to the skedastic function are predetermined with respect to time t information.

$$\mathbb{E}[\zeta_{i,j,t+1}|\mathcal{I}_t] = 0$$

The empirical setup uses GMM and these timing assumptions, in addition to the structure of the model, provide population moment conditions that can be used to implement my strategy.

2.1.3 Implementation

I have now characterized all that is required about the environment. Using these assumptions I can now discuss the empirical implementation. The strategy is standard in the literature, but I think it is useful to discuss my extension to the more general Markov Process. This is because Olley and Pakes (1996), Doraszelski and Jaumandreu (2013) and Buettner (2004) are all interested in similar identification issues. None of these papers use my strategy of inverting the materials function. Olley and Pakes (1996) and Buettner (2004) impose additional restrictions to the stochastic process in order to guarantee invertibility of the investment policy function. I have not discussed the investment policy function as it is not necessary for my strategy. Doraszelski and Jaumandreu

[11] I could have instead put restrictions on the gross production function and built this assumption up from the static profit function.

(2013) require variation in input and output prices to identify their production function. Inverting materials generalizes their approach and will I think be useful in future applied work. In particular, the use of representative datasets, such as the LBD, is important when analyzing firm dynamics. This is because firm size distributions have very fat tails.[12]

The estimation strategy proceeds in three stages. In the first stage I estimate:

$$y_{i,j,t} = \iota_j + \tau_t + h(\mathbf{m}_{i,j,t}, \mathbf{x}_{i,j,t}) + \beta_x \mathbf{x}_{i,j,t} + \varepsilon_{i,j,t}$$

Which can be simplified to:

$$y_{i,j,t} = \iota_j + \tau_t + H(\mathbf{m}_{i,j,t}, \mathbf{x}_{i,j,t}; \beta_x) + \varepsilon_{i,j,t}$$

The function $H(\mathbf{m}_{i,j,t}, \mathbf{x}_{i,j,t}; \beta_x)$ is of unknown form and is approximated by a second-order polynomial. I have been explicit that this function subsumes the production function parameters.

In the second stage I make use of the threshold exit policy to find an exit probability for the firm. The probability of a firm exiting will depend on materials and dynamic inputs in a complex way. I estimate a probit model using the 2nd order polynomial in materials, labour and capital from stage 1. The exit dummy is used on the left hand side. I use this probit estimation to recover exit probabilities, $\hat{P}(\mathbf{m}, \mathbf{x})$.[13]

In the final stage I estimate the production function coefficients, $\hat{\beta}_x$. I do this by first recognizing that I can recover productivity from stage 1 as follows:

$$\hat{z}_{i,j,t} = \hat{H}_{i,j,t} - \hat{\iota}_j - \hat{\tau}_t - \hat{\beta}_x \mathbf{x}_{i,j,t}$$

I can do likewise for $\hat{z}_{i,j,t+1}$. It is important to note that these values for productivity are conditional on the choice of estimate, $\hat{\beta}_x$. Now I can use the controlled markov process to run the following regression:

$$z_{i,j,t+1}(\beta_x) = g(\hat{z}_{i,j,t}, r_{i,j,t}, \mathbf{w}_{i,j,t}) + \sigma(z_{i,,t}, r_{i,j,t}, \mathbf{w}_{i,j,t})\xi_{i,j,t+1}$$

I do this for different specifications of $g(\bullet)$. I can also store the residuals to estimate the skedastic function:

$$\hat{\xi}^2_{i,j,t+1} = \sigma^2(z_{i,j,t}, r_{i,j,t}, \mathbf{w}_{i,j,t}) + \zeta_{i,j,t+1}$$

The residuals from both regressions are orthogonal to any information not known by the firm at time $t+1$. This guides the choice of instruments. As I now have that:[14]

$$\mathbb{E}[\varepsilon_{i,j,t+1} + \sigma(z_{i,j,t}, r_{i,j,t}, \mathbf{w}_{i,j,t})\xi_{i,j,t+1} | \mathcal{I}_{i,j,t}] = 0$$

[12] I realize that this point may weaken my paper as I am using Compustat data. However, the Compustat data captures precisely the firms which are in the tails. It neglects to capture small, non-publicly traded companies. From the point of view of R&D , these are less interesting.

[13] When I "control" for exit I include the exit probabilities on the right hand side of my TFP process. The resulting results I present control for exit, and are appropriate for firms for which $\hat{P} = 0$, that is continuing firms. These are the appropriate stochastic process estimates to feed into my quantitative model in Chapter 3.

[14] I am trying to find a GMM based paper to cite defending this orthogonality condition. If not, I need to think a bit more rigorously about whether the skedastic function changes the interpretation. I actually do not use the residuals from the skedastic function at this point. I just estimate it in addition to the production function.

Notice that this strategy reduces the minimization step to finding the coefficients, β_x, since the hard work is done by these two regressions. As well, notice that the only assumptions that are necessary for this procedure to be valid are the two I have numbered. In particular, conditional heteroskedasticity does not affect the estimation procedure for the production function coefficients provided the factors that influence the volatility are predetermined. This is what allows me to run the auxiliary regression to identify the skedastic function.

I use the bootstrap to compute the standard errors. The re-sampling procedure is as follows: each bootstrap sample includes the same number of firms as the original sample. The firm-year observations may be slightly different as I select all years a firm is active if a firm is sampled. I use 299 bootsrap replications when I estimate the production function.

2.2 Context Within the Literature

At this point I will discuss why this approach is useful and how it differs from the literature. First, as I mentioned above, the problem of dealing with stochastic accumulation of knowledge has been addressed in two ways in the literature. Doraszelski and Jaumandreu (2013) rely on variation in wages and prices while the Olley and Pakes (1996) approach requires inversion of the investment function. This becomes much more challenging once stochastic accumulation is introduced. This is because the capital policy function must be strictly monotonic in productivity for the approach to be valid. When endogenous R&D spending influences the future distribution of productivity in non-linear ways this result is impossible to prove. Buettner (2004) avoids this by restricting the way that R&D spending can influence productivity dynamics. This is unsatisfying as the restrictions are rejected in several industries in his paper.

Specifically, he requires that the distribution of productivity next period for higher levels of R&D must stochastically dominate the distribution for lower levels. This allows him to invert the investment demand function. When he then examines the productivity distributions implied by his estimates he finds that in many industries the stochastic dominance assumption is violated. This is not surprising in the context of my paper. I will show that R&D affects productivity dynamics in non-linear ways, violating the restriction in this paper.

I now compare my approach to Doraszelski and Jaumandreu (2013). I have already mentioned how my approach applies to data sets which do not feature variation in input prices. This is the first contribution: generalizing the approach and estimating a controlled markov process on the Compustat data.

My second contribution is to explicitly model the life cycle of the firm through the productivity process. There is a lot of research being done on the life cycle of firms. In particular, young firms grow faster, have more uncertaint growth rates and are more likely to exit. Pugsley et al. (2018) inspect the covariance matrix of employment while Decker et al. (2014) show the percentiles of the employment growth distribution by age. Both of these approaches use employment to characterize the post-entry dynamics of firms. I use TFP, which can only be done in the context of estimates for the production function. This is why I consider an age dependent TFP process. In principle including R&D and other types of intangible capital as drivers of TFP could mitigate the importance of age. I find that these richer processs still include strong age effects. This has implications for how I interpret theories of firm growth.[15]

My next contribution is to explicitly model uncertainty through a skedastic function. Given the importance of uncertainty in the life cycle of firms and the uncertainty inherent in R&D , I think that this is a novel contribution. Doraszelski and Jaumandreu (2013) show that industries with more R&D have higher uncertainty estimates in their stochastic process.[16]

The aim of my approach is to first model TFP dynamics in the richest possible manner. Once I have identified the essential elements of such a process I will feed it into a firm dynamics model a la Hopenhayn and Rogerson (1993). The inclusion of R&D brings my work close to the endogenous growth literature. In fact, state of the art endogenous growth models have rich firm heterogeneity, with different types of firms conducting different types of R&D .[17]

I take a complementary, but different, approach to modeling R&D dynamics. First, I do not distinguish between types of R&D . Instead I allow interactions between productivity, age and R&D which are non-linear. This means, for example, that older firms could be more or less productive at R&D . My data identifies this. I also allow for interactions between productivity (size) and R&D in a similar way.

I do not focus on a balanced growth path. I attempt to identify spillovers which have level effects on aggregate TFP. My hope is that this disciplines my policy analysis.[18] The spillovers I identify motivate a role for policy. There are many policies which could be relevant. Subsidies or taxes to R&D spending are the obvious example. However, making these subsidies size and age dependent is of direct relevance in my paper. **These are the types of policies I will consider in Chapter 4 of this thesis. The next step in my empirical project is to measure "closeness" in R&D space using**

[15] I am going to include a general discussion of theories of firm growth in my first chapter. These are: Intangible Capital Accumulation (R&D and Advertising), financing constraints, tangible capital accumulation and pure age effects (such as learning). So far, my results suggest that pure age effects remain after I condition on R&D and tangible capital accumulation. I can write a secondary paper which conditions on advertising and financing constraints. In essence my methodology provides a nice way to test multiple models of firm growth using one unifying framework. **This is a distinct paper I could work on that I think would be a contribution to the literature.**

[16] Explicitly, if the process is:

$$z' = \rho z + \gamma r + \sigma \varepsilon'$$

They find that σ is higher in industries which are more R&D intensive. This is suggestive of the mechanism I have in mind, but I model it more directly.

[17] Akcigit and Kerr (2018) is a cutting edge example. Large firms do more internal R&D on existing product lines while small firms and entrants do external R&D , looking for new product lines. This heterogeneity in R&D provides a mechanism whereby small firms grow faster and contribute disproportionately to creative destruction. They abstract from firm age.

[18] A complicated model is not required to argue that any policy which affects growth rates will have large effects. The problem is that variation in aggregate R&D does not seem to be associated with variation in growth rates. **Think seriously about this critique of growth models if I am going to take such a strong stance.**

patent citation data and come up with a "closeness weighted" measure of R&D . This will be used to identify spillovers.

2.3 Introduction to the Models

In this section I introduce the various models of TFP that I use. The general notation is:

$$z' = g(z, r, \mathbf{w}) + \sigma(z, r, \mathbf{w})\xi'$$
$$\hat{\xi}'^2 = \sigma^2(z, r, \mathbf{w}) + \zeta'$$

Where z is productivity, r is R&D expenditures and $\mathbf{w}$ is a vector that contains age, aggregate productivity, etc.

The firms dynamic program leads to an Euler Equation for firms which do positive R&D and continue operation:

$$\frac{\partial \Psi_r(r)}{\partial r} = \beta \int \frac{\partial \pi(z', \mathbf{x}')}{\partial z'} dG(z'|z, r, \mathbf{w})$$

The interpretation of this equation is that the marginal costs of R&D are equated to the marginal benefits. Given fixed costs, some firms will choose the corner solution of 0 R&D . As well, some firms will exit. The Euler equation represents the intensive margin decision of "how much" R&D to do while the extensive margin is related to the fixed costs.

The simplest model of R&D would involve only fixed costs and would model it as a discrete choice. Firms either do R&D and have a stochastic process:

$$z' = \gamma_0^r + \rho^r z + \sigma^r \xi'$$

or they do not do not and have the process:

$$z' = \gamma_0 + \rho z + \sigma \xi'$$

This basic model of R&D actually has a lot of content. The discrete decision to do R&D could lead to a higher mean ($\gamma_0^r > 0$), more or less persistence ($\rho^r \neq \rho$) and more or less uncertainty ($\sigma^r \neq \sigma$).

Next, I introduce a simple model of R&D with an intensive margin decision which affects both the mean and variance of the TFP process. I model this as:

$$z' = \begin{cases} \gamma_0^r + \rho^r z + \gamma_r r + \xi' & r > 0 \\ \rho z + \xi' & r = 0 \end{cases}$$

$$\hat{\xi}'^2 = \begin{cases} \sigma_{0,r}^2 + \sigma_r^2 r & r > 0 \\ \sigma_0^2 & r = 0 \end{cases}$$

Now the level of R&D impacts the evolution of productivity. Significantly, this effect works through both the first and second moments of productivity.

As a last step, I allow interactions between R&D and age as well as R&D and productivity. These interactions are one way of modelling the heterogeneity in the behaviour of R&D . In particular, there are reasons to believe that young firms do R&D projects which are inherently different from more mature firms. For example, the endogenous growth model of Aghion and Howitt (1992) has the property that all new innovations are introduced by entrants. This is due to the replacement effect introduced by Arrow (1962). This can be relaxed as in Klette and Kortum (2004) where both entrants and incumbents do R&D . However, incumbents are less likely to introduce new product lines. Furthermore, the mechanism of internal vs. external innovation in Akcigit and Kerr (2018) implies that firms with higher attained productivity will innovate differently from those with lower productivity.[19] To be concrete, I estimate variations on the following stochastic process:

$$z' = \gamma_a a + \gamma_{aa} a^2 + \begin{cases} \gamma_0^r + \rho^r z + \gamma_r r + \gamma_{rz} r z + \gamma_{ra} r a + \xi' & r > 0 \\ \rho z + \xi' & r = 0 \end{cases}$$

Notice how I have included two life cycle terms. This is to test whether there are strong life cycle effects independent of the R&D effects as well as how R&D interacts with age. In particular, the partial derivative of productivity with respect to age is:

$$\mathbb{E}[\frac{\partial z'}{\partial a}|z, a, r] = \gamma_a + 2\gamma_{aa} a + \mathbb{I}(r > 0)\gamma_{ra} r$$

This can inform whether average productivity grows over the life cycle and how it interacts with R&D . The partial derivative of productivity growth with respect to age is:

$$\mathbb{E}[\frac{\partial \Delta z'}{\partial a}|z, a, r] = 2\gamma_{aa} + \mathbb{I}(r, r_{-1} > 0)\gamma_{ra} \Delta r$$

Given that firm growth is declining over the life cycle I include the quadratic life cycle term to capture such dynamics.

It is also informative to consider the marginal effects of increasing the amount of R&D expenditures:

$$\mathbb{E}[\frac{\partial z'}{\partial r}|z, a, r] = \mathbb{I}(r > 0)[\gamma_r + \gamma_{ra} a + \gamma_{rz} z]$$

This makes explicit how the interaction terms inform the interpretation of the benefits of R&D .

3 Sample Selection and Summary Statistics

This section of the paper introduces the dataset which will be used to carry out the procedure outlined in Section 2. In particular, Section 3.1 discusses how the relevant

[19] The micro-economics of R&D has proposed many dichotomies in "types" of research. Internal vs. external, process vs. product, drastic vs. non-drastic. My approach is to abstract from these difficult to measure distinctions and allow the level of R&D to interact with measurables such as age and productivity. There are merits to both approaches but I think given the complexity of the R&D process my level of abstraction is useful in a macro model.

variables are constructed using Compustat and the choices I make regarding sample selection. In Section 3.2 I analyse the firm life cycle using data on employment. This is in the interests of comparability with a recent literature on the empirics of firm growth using Longitudinal Business Database (LBD) data. Qualitatively I find a similar life cycle in the data on publicly listed companies. Section 3.3 contains a detailed discussion of the pros and cons of my dataset.

3.1 Data and Sample Selection

The principal data source used in this paper is the Standard and Poor's Compustat database. This data is suited to this study as it covers the substantially long period of 1950 to 2015. As well, since the data are constructed from financial statements of publicly traded companies, there are well defined measures of sales, R&D expenditures, the ownership structure of the firm as well as advertising and sales expense. This will be extremely useful in comparing different models of firm growth.

Appendix 1 gives a detailed summary of the variable construction and sample selection procedure. I will briefly discuss how age, materials, value added, capital stock and R&D expenditures are measured as they are the key variables in the study. I construct age as years since listing and control for mergers and acquisitions. In this I attempt to mimic the "organic age" measures constructed using LBD data. If a firm undergoes merger activity which does not change its sales by a large amount, I consider it to be the same firm. If it undergoes larger merger activity then I drop it from the sample. As a consequence, this makes exits noisier in my data than in the LBD as they represent sample attrition due to mergers as well as natural exit. My focus on the life cycle of firms motivates this choice. I want to be sure to have clean measures of firm ageing at the expense of noisy exits.

Materials and value added are constructed together as they are intricately related. Gross revenue can be measured as net sales. In order to turn this into a measure of value added I must subtract some materials measure. Compustat measures operating income before depreciation. This is defined as sales minus cost of goods sold and selling, general and administrative expenses. The R&D expense is included in selling, general and administrative expenses while labour expense is included in cost of goods sold. I define materials as:

$$\text{M} = \text{COGS} + \text{SGA} - \text{Labour Expense} - \text{R\&D Expense}$$

Where the R&D expense is given in the data and is sometimes 0. Labour expense is trickier. There are measures of staff expense in the Compustat data but they are sparsely populated. The number of employees is available for almost all observations. I use data from the Social Security Administration (SSA) on average wages to construct wage expense as:

$$\text{Labour Expense} = \bar{w}N$$

The fact that there will be deviations from average wage will introduce some measurement error to my results.

My target sample is the universe of non-subsistence entrepreneurs. Given that the decision to go public signals an intention to grow the company, the Compustat data are a starting point for capturing this sample. My data limit me from measuring non-subsistence entrepreneurs which have not listed yet, or which never list. The sample selection choices I make are all in the interests of data quality. I throw out firms which existed in the sample before 1950 as I cannot measure their age. I throw out public utilities and financial firms. I exclude extreme outliers. The full procedure is described in **Appendix 1.**

One potential weakness of the dataset is its lack of coverage. Compustat data includes all publicly traded firms. While this represents a small fraction of the entire universe of firms, it does include a substantial fraction of economic activity. In particular, 29% of private US employment is at firms in the Compustat sample. **R&D and advertising expenditures would be good to know as well here, given the focus of this study.** The concern is that results from this non-representative sample cannot be applied to the entire universe of firms. In particular, the subset of firms which choose to go public are not random. I will address the ways in which I account for this throughout.[20]

Table 1: Summary Statistics

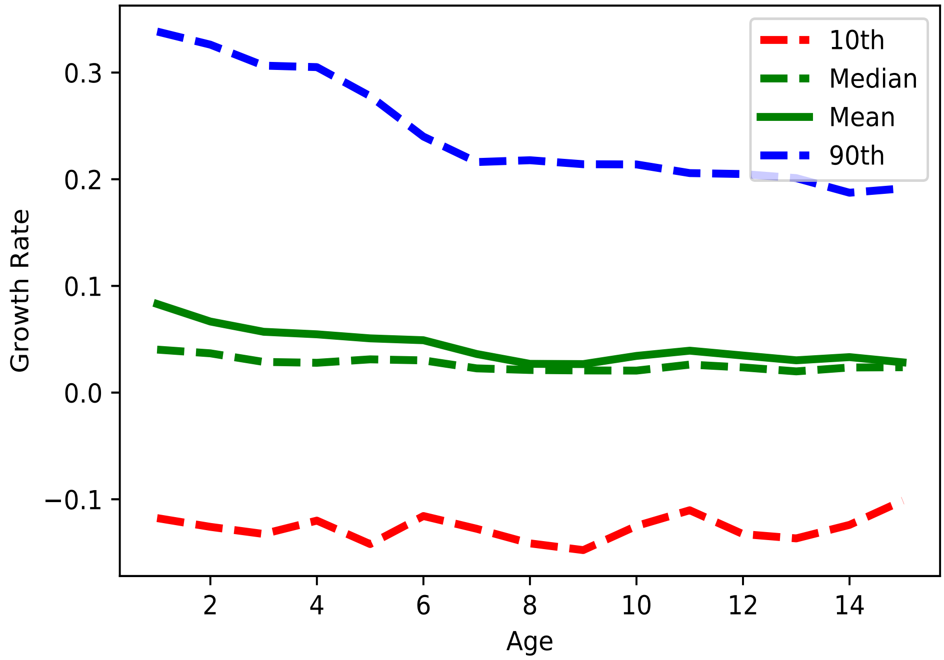
	Mean	Median	10th Perc.	90th Perc.	St. Dev
Value Added	229.42	70.36	4.35	633.77	419.96
Capital	373.36	54.54	3.09	933.95	963.95
Employment	3.15	0.97	0.08	8.28	5.90
R&D	7.85	0.00	0.00	21.19	22.78
Materials	528.33	124.59	9.65	1305.23	1481.01
Investment	38.14	6.76	0.27	102.68	86.50
Age	10.68	8.00	1.00	25.00	10.27
Observations	83680				

*Value added, capital, R&D , materials and investment are measured in millions of constant 2012 US dollars. Employment is meaasured in thousands of employees and age is measured in years.

As a result of the concerns raised in the previous paragraph, the firms in my dataset are on average much bigger than the entire universe of firms. They do more R&D , they exit less, they have less uncertain life cycles and they employ more than privately held firms. These results are well known and are summarized in **Table 1.** For instance, average value added is approximately 230 million, while the 90th percentile is approximately 634 million dollars. These numbers are very large. Compustat features some of the worlds

[20]One way to think about the sample selection is as follows. In my quantitative model I will include a publicly traded sector, for which I have estimated the production function and stochastic process for productivity. I will also include a subsistence entrepreneurial sector and a young but not yet incorporated sector. Every period a young unincorporated firm will incorporate with some probability. The subsistence entrepreneurs will not make efforts to grow. The underlying assumption is that these two sets of firms are distinct. I motivate this simplification by the findings in Hurst and Pugsley (2011) which convincingly show that there is a large subset of firms which do not aspire to become large.

Figure 1: The Life Cycle of Firm Growth



*The weighted percentiles of the employment growth distribution by age. I use average employment as weights. The measure of employment growth is the average growth measure.

largest companies. Furthermore, average employment is just over 3,000 employees, while the 90th percentile is almost 6,000 employees.

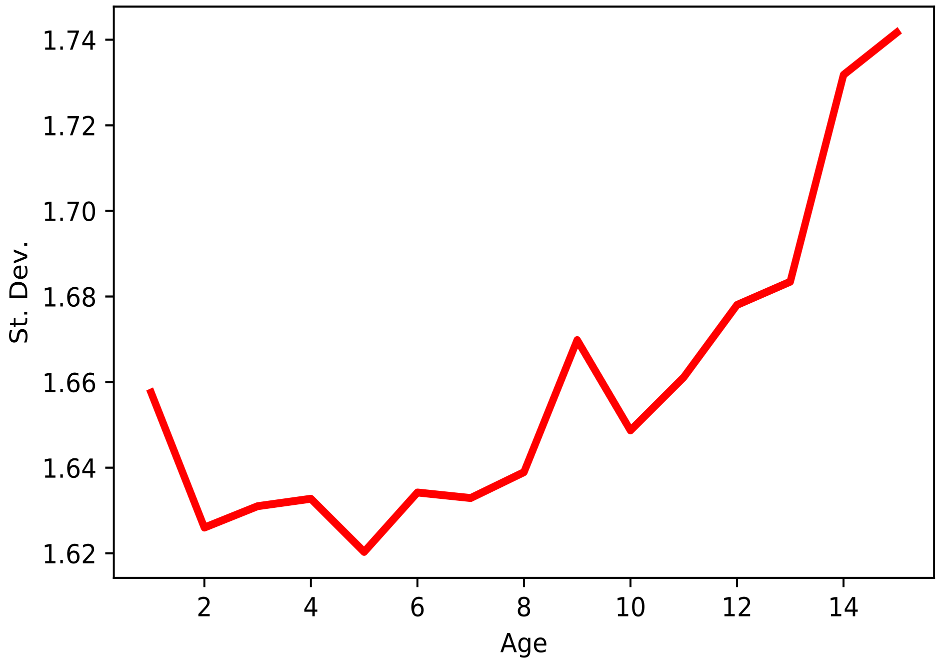
A striking feature of the data is that all variables are highly right skewed. The median value added, capital, employment, R&D and materials measures are all much lower than their means. This is a common feature of all firm level datasets and can be seen in the distributions of **Figure A1.**

3.2 The Firm Life Cycle

This paper is motivated by a recent literature on firm life cycle dynamics which mainly uses Census Bureau data. Since I want to relate to this literature, I examine the Compustat data to see if they follow roughly the same set of stylized facts. As I discuss in this section I find some similarities and some differences.

A well known piece of evidence on the firm life cycle is the shape of the employment growth rate distribution as firms age. Decker et al. (2014) emphasize that young firms have employment growth distributions which have higher variances, higher skewness and higher means than more mature firms. They also find that exit rates decline over the life cycle. These striking facts motivate my entire thesis. As such, **Figure 1** replicates

Figure 2: The Standard Deviation of Employment Over the Life Cycle



*I regress log employment on 2 digit SIC dummies as well as a year dummy. I find the standard deviation of the residuals by age.

their figure using my dataset.

I use average employment growth as my measure and construct the employment weighted 10th, 50th and 90th percentiles as well as the mean growth rate by age.[21] The first thing I note is that the Compustat data, using age since listing, qualitatively matches the facts emphasized in Decker et al. (2014). The mean, variance and skewness of growth rates decline over the life cycle. The second point to note is that the level of uncertainty is lower for these publicly traded firms. The 90th percentile for an age 1 firm in my data is approximately 33% while, it is about 90% in the LBD data.

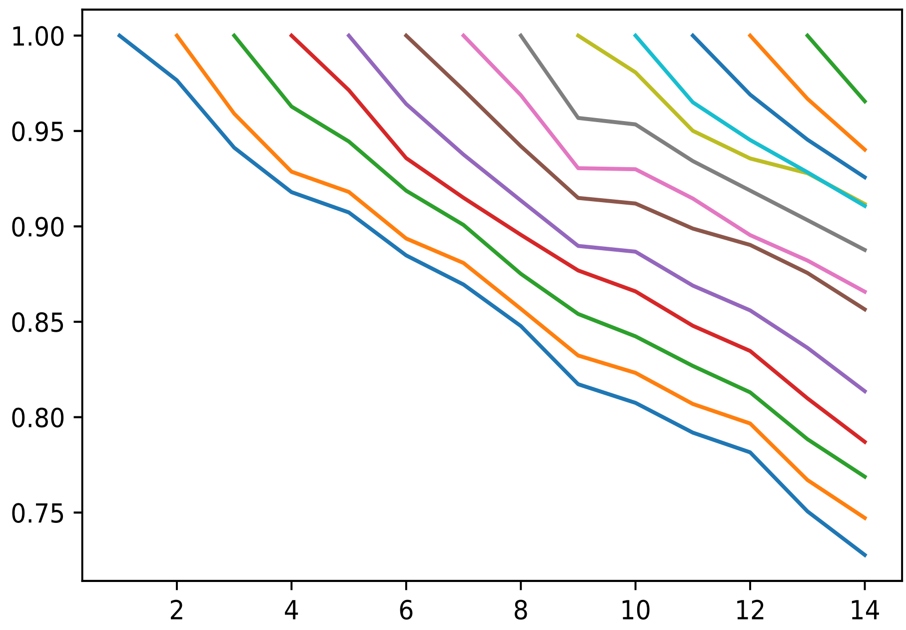
One possible reason for this is that publicly traded firms do not list right away. If this is the case, then I am picking up the life cycle dynamics of firms starting at a more mature point in the life cycle. This would explain the lower levels of dispersion. An interesting and related issue is whether listing is an event which dramatically alters the

[21]In particular, average growth is:

$$g = 0.5\frac{N_{i,j,t+1} - N_{i,j,t}}{N_{i,j,t+1} + N_{i,j,t}}$$

It has the property that it is bounded between -2 and 2, with -2 denoting an exit and +2 denoting entry. As well, regression to the mean effects are mitigated. I make these choices in the interests of comparability with data that represents the entire universe of firms.

Figure 3: The Correlation of Employment Over the Life Cycle



*I regress log employment on 2 digit SIC dummies as well as a year dummy. I find the correlation between employment at age a and age $a + j$.

life cycle of the firm. There are reasons to believe this may be the case. The ownership structure of the firm could have an effect on its risk properties. For instance, shareholders may dislike risk and the firms management may therefore behave more conservatively. With these caveats in mind, I progress with the comfort that the qualitative properties of the firm life cycle are well established in my dataset.[22]

The percentiles of the growth rate distribution convey the "up or out" dynamic in a striking manner. However, there are other measures of the firm life cycle which may be more informative. In recent work Pugsley et al. (2018) characterize the full covariance matrix of log employment. In particular, they document that the standard deviation of log employment grows over the life cycle. This means that the dispersion of employment is higher for more mature firms. They also point out that the covariance between log employment across ages declines with age, but never diminishes to 0. They convincingly use this as evidence of very persistent differences in firm type.

Figure 2 plots the standard deviation of log employment over the life cycle using the methodology of Pugsley et al. (2018).[23] After an initial decline in the dispersion of

[22]I have already noted that exits are noiser in my dataset. **Figure A2** shows the exit rates over the life cycle.

[23]In particular, I regress log growth on a set of time and industry dummies and store the residual. I

employment, the pattern of rising dispersion over the life cycle is present. The level of this dispersion is much higher in the Compustat data than the LBD. For instance, the standard deviations range from 1-1.4 log points in the LBD, while they range from 1.64 to 1.74 in my data. This reflects the fact that there are a disproportionate number of large, growing firms in my sample and fewer small subsistence entrepreneurs. The initial decline may also reflect the fact that listing is a significant event in a firms life cycle.

Figure 3 plots the auto-correlations of log employment over the life cycle. I am able to replicate the result from Pugsley et al. (2018) that these auto-correlations do not diminish to 0. This is evidence of fixed types for firms, which do not disappear over the life cycle.

3.3 Discussion

In this section I have described my data creation and sample selection procedure and compared the sample to known facts on the firm life cycle using more representative samples. This comparison informs my interpretation of my results. In this section I will defend my use of Compustat for this project and talk about how knowledge of its shortcomings qualifies the conclusions I will draw.[24]

There are certain aspects of the firm life cycle that can be uniquely examined with Compustat data. In particular, the dataset includes advertising and R&D expenditures. This allows me to test two distinct theories of intangible capital accumulation in my framework. The data also include measures of financing constraints as I can view the debt and ownership structure of the firms. For instance, the amount of equity and debt. I also can view market capitalization as an estimate of the value of the firm. To the extent that I can incorporate these in my analysis, the Compustat data are useful.

The NBER Patent Database has been merged with Compustat.[25] This is very useful for my identification of spillovers. The reason is that patent information can be used to construct measures of "closeness" in technology space between firms. Bloom et al. (2013) estimate a model of knowledge capital with closeness weighted measures of R&D to identify spillovers. I will use their measure of aggregate R&D but in the context of my framework. In particular, I can test whether aggregate R&D affects the evolution of firm level productivity:

$$z' = g(z, r, \mathcal{R}) + \sigma(z, r, \mathcal{R})\xi'$$

To the extent which aggregate R&D , $\mathcal{R}$, influences individual firm TFP, there are spillovers. These spillovers motivate a role for policy in either subsidizing or taxing

then calculate the standard deviation of this residual for different ages.

[24]Currently there is a data liberation effort occurring in Canada which will allow me to repeat much of my analysis on a more representative sample. The Productivity Partnership, along with StatsCan are hoping to make firm level data available to researchers in the RDCs by next summer. However, this data is not linked to patent statistics yet. In the US some recent studies have linked the LBD with R&D expenditures and patent statistics. However, advertising, firm value, etc. are not available yet.

[25]See Hall et al. (2001) for a discussion of this dataset.

R&D expenditures. My empirical procedure provides a way for the data to discipline the magnitude of such spillovers.

The downside to using the Compustat data is of course the non-representativeness of the sample and the concern that age is mis-measured. I argue that the rich information on firm performance, R&D , patent citations, advertising and financing constraints in addition to the qualitative similarity to more representative samples justifies using this data. There is much that can be learned from the dataset if appropriate qualifications are made.

A main theme of my project is firm level uncertainty over the life cycle. To the extent that Compustat under represents the most uncertain firms, all these results can be considered conservative. For instance, when examining the interaction between uncertainty, R&D and selection I will be using estimates from a sample which is on average less uncertain than the entire population of firms. Any quantitative effects which my model delivers could be considered lower bounds on the magnitude of such effects for the macroeconomy.

4 Main Results

In this section of the paper I present my main results regarding productivity and R&D at the firm level. Section 4.1 presents the production function elasticities under different modelling assumptions. Section 4.2 presents estimates for the TFP process. In Section 4.2.1 I present estimates of the conditional mean, $g(z, r, \mathbf{w})$ while Section 4.2.2 presents estimates for different specifications of $\sigma^2(z, r, \mathbf{w})$. Section 4.3 discusses productivity over the firm life cycle and Section 4.4 discusses the implications of my findings for the knowledge capital literature.

4.1 The Production Function

Let's begin by examining the production function elasticities. For all of these results I will show estimates for 5 different models of TFP growth. The first is an AR(1) process which is homoskedastic. This process has a long tradition in the macro literature on firm dynamics and is the driving force in Hopenhayn and Rogerson (1993). My estimation strategy is of course quite different from much of this literature so there is no reason ex-ante to guess that I should find the same levels of volatility and persistence as Castro et al. (2015), for example. In fact, I find greater volatility and greater persistence than them.

One thing that is clear from Table 2 is that the production function estimates are very stable across specifications of the TFP process. The only model which is notably different is the OLS specification, which isn't surprising due to bias of the estimates. The stability of the production function estimates is useful for examining TFP over the life cycle. It appears the specifics of the TFP model used in the estimation strategy are of secondary importance in recovering the TFP residual.

Table 2: Production Function Estimates

	OLS	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
β_k	0.2571	0.1533	0.1512	0.1536	0.1505	0.1509
	(0.0083)	(0.0265)	(0.0233)	(0.0275)	(0.0240)	(0.0241)
β_n	0.6855	0.8839	0.8819	0.8533	0.8532	0.8532
	(0.0106)	(0.0441)	(0.0349)	(0.0542)	(0.0415)	(0.0349)

*Production function elasticities for different specifications of the stochastic process. Bootstrap standard errors in parentheses. In this table I have controlled for exit.

In these results I have controlled for endogenous exit. It turns out that this does have a sizeable influence on the parameters of the production function. In Section 6 I present estimates which do not control for exit. These estimates are very similar to the OLS elasticities.

There is general consensus in the production function estimation literature that controlling for exit is not of first order importance for estimating these elasticities. My findings seem to go against this belief. There are two potential reasons. First, it could be because I am using data on materials which introduces extra noise. This is because I have assumed all employees are paid the average wage. I try to ameliorate concerns around this by estimating my production function on a restricted sample, which uses the wage and salary expense available in compustat. This has the downside of further restricting the sample.

A second reason for the difference could be my estimation strategy. I am including more terms in the TFP process than is standard. This could fundamentally change the nature of identification and could lead to different conclusions on the relative importance of controlling for endogenous exit. **I have not yet done the robustness exercise on the materials measure.**

4.2 The Stochastic Process

In this section I analyse the estimates of the TFP process for different specifications. I first look at the "first moment" effects which are summarized in Table 3. That is, what is the function $g(z, r, \mathbf{w})$ which gives the conditional mean of next periods productivity. In the second subsection I give results for the volatility effects in the form of the skedastic function, $\sigma^2(z, r, \mathbf{w})$. These results are given in Table 4.

4.2.1 First Order Effects

In order to interpret these results, I first note that all of these models are restrictions on the following general model:

$$z' = \gamma_a a + \gamma_{aa} a^2 + \begin{cases} \gamma_0 + \rho z \\ \gamma_0^r + \rho^r z + \gamma_r r + \gamma_{rz} rz + \gamma_{ra} ra \end{cases} + \sigma(z, r, \mathbf{w})\xi'$$

Table 3: Stochastic Process Estimates

	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
γ_0	-0.1257^{**}	-0.1283^{**}	-0.0998	-0.1010^{**}	-0.1005^{**}
	(0.0440)	(0.0409)	(0.0581)	(0.0458)	(0.0384)
γ_a	-	0.0061^{**}	-	0.0040	0.0039^{*}
	-	(0.0029)	-	(0.0036)	(0.0029)
γ_{aa}	-	-0.0010	-	-0.0007	-0.0007
	-	(0.0006)	-	(0.0007)	(0.0006)
ρ	0.9247^{***}	0.9266^{***}	0.9399^{***}	0.9412^{***}	0.9413^{***}
	(0.0191)	(0.0170)	(0.0237)	(0.0192)	(0.0165)
γ_0^r	-	-	-0.1752	-0.1752^{**}	-0.1787^{**}
	-	-	(0.0815)	(0.0597)	(0.0472)
ρ^r	-	-	0.8188^{***}	0.8219^{***}	0.8144^{***}
	-	-	(0.0673)	(0.0434)	(0.0325)
γ_r^r	-	-	-	0.0007	0.0036
	-	-	-	(0.0010)	(0.0035)
γ_{rz}^r	-	-	-	-	0.0061
	-	-	-	-	(0.0061)
γ_{ra}^r	-	-	-	-	0.0002
	-	-	-	-	(0.0003)

*Stochastic process estimates for various specifications. These correspond to the elasticities in Table 2. Stars denote level of significance. * ⇒ Significant at 10% level, ** ⇒ Significant at 5% level. *** ⇒ Significant at 1% level.

For instance, the AR(1) model restricts $\rho = \rho^r$ and $\gamma_0 = \gamma_0^r$ while $\gamma_a = \gamma_{aa} = \gamma_r^r = \gamma_{rz}^r = \gamma_{ra}^r = 0$. This is useful as I could do more formal statistical testing in the form of F-tests for model selection. For now I note that in the most general model all terms are significant at least at the 10% level. Furthermore, as I describe, the interpretation of the model changes significantly when the full interaction terms are included. For this reason, the most general model is my preferred specification.

Regardless of the specification, there are some results which are clearly evident. First, productivity is very persistent with $\rho \approx 0.94$ for non-performers and $\rho^r \approx 0.82$ for R&D performers. This has interesting implications for the extensive margin decision to do R&D . High productivity firms would be less likely to do R&D as they would prefer a higher persistence. Firms which have not accumulated as much productivity would be more likely to engage in R&D . The average productivity for the productivity performing firms is lower, perhaps because many R&D projects do not work out.

Second, the pure life cycle effects are not significant in most cases and are not quantitatively large, even when significant. When interpreting the magnitudes of γ_a and γ_{aa} it is important to remember that I have included log age in this specification. For instance, the elasticity of TFP with respect to age for a non-R&D performer is:

$$\gamma_a + 2\gamma_{aa} a$$

This means that age can effect TFP in non-linear ways over the life cycle. The point estimates suggest that the profile of TFP as a firm ages is concave. This is because, $\gamma_{aa} < 0$. The elasticity of productivity growth with respect to age is $\gamma_a \Delta a + \gamma_{aa} \Delta a^2$. Where Δa represents a percentage change in age. The percentage change in age of a year is declining over the life cycle. In the full specification I find that $\gamma_a \approx 0.004$ and $\gamma_{aa} \approx -0.0007$ implying an increasing, concave profile for productivity over the life cycle. Bear in mind that this is revenue productivity so that my measure confounds both supply and demand side effects.

In order to comprehend the magnitude of these life cycle terms, consider the average firm of age 1 and 5. The life cycle component implies that the age 5 firm will be 0.5% more productive than the age 1 firm. Productivity growth of the age 1 firm will be 0.24% while productivity growth for the age 5 firm will be 0.03%. Of course this is just due to the life cycle profile. I have not discussed the interaction between age and R&D spending yet. At this point it appears that the age profile is not quantitatively important in my model. This is interesting, as much focus has been placed on age the relationship between firm age and growth in Decker et al. (2014) and the surrounding literature. My results suggest that it is not age per se that matter for productivity growth. Including a measure of intangible capital essentially kills off what little pure age effects there were.

I now interpret the intensive margin of R&D , γ_r. I find that $\gamma_r \approx 0$. This means that doing more R&D does not have a significant effect on the average productivity next period. This does not change once I allow for interactions between R&D and age or productivity. In fact the intercept terms are not statistically different from each other either. This suggests that for the function $g(\bullet)$ the extensive margin choice of whether to do R&D is most important. This has an impact on the persistence parameter, but no other properties of the rich non-linear function I have estimated.

My analysis of these specifications suggests that a process of the form:

$$z' = \mathbb{I}(r=0)\rho z + \mathbb{I}(r>0)\rho^r z + \sigma(r,z,a)\xi'$$

Will represent the salient features of the data.

4.2.2 The Skedastic Function

One of the key contributions of this paper is the systematic analysis of the conditional heteroskedasticity with respect to R&D spending. In fact, my results on the skedastic function are much stronger than the results I find on the conditional mean function, $g(\bullet)$.

I start with the surprising finding that the average variance of an R&D performing firm is slightly lower than a non-R&D performing firm. That is, $\sigma_0^2 \approx 0.67 > 0.58 \approx \sigma_{0,r}^2$. This finding masks a lot of heterogeneity over the life cycle and with respect to R&D . First, I find that the elasticity of the variance with respect to age is -0.072. A firm which ages by 1% sees its variance decline by 0.07%.

Next, I find that the elasticity of R&D spending with respect to the variance of shocks is 0.13%. This is what I find in my full model with interactions. In the model without interactions I find that the elasticity is positive, but an order of magnitude lower. This once again points to the potential importance of these non-linearities.

The interaction between volatility and already attained productivity is positive. This means that larger firms are performing R&D which is more volatile. This is at odds with the internal vs. external innovation mechanism in Akcigit and Kerr (2018).

Lastly, I find that the uncertainty of R&D efforts is declining over the life cycle. The interaction between age and R&D spending is negative and consistent. This suggests that as firms age they perform less risky R&D .

4.2.3 In Summary

My methodology has uncovered a complex relationship between R&D and TFP evolution which is highly uncertain. The amount of R&D that firms perform impacts both the level and the volatility of TFP next period. It does so in ways which change over the life cycle and with the success of past R&D projects. In some sense, these results are not surprising. The literature on R&D has proposed many different classifications for "types" of research. These types have unique implications for the first and second moments of future productivity. The fact that the literature has identified so many mechanisms speaks to the inherent heterogeneity in the process of turning exploration and R&D effort into useful gains to productivity. [26] One benefit of my approach is to abstract from some of the micro-economic heterogeneity and identify non-linearities which are related to easily measurable characteristics of the firm.

In particular, it is perhaps surprising that R&D works primarily through the volatility channel and through persistence. As I will show in Chapter 3 this has interesting

[26] I include both new products and processes here since my measures all confound the two.

Table 4: Skedastic Function Estimates

	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
σ_0	0.4839***	0.6491***	0.5226***	0.5954***	0.6672***
	(0.0081)	(0.0172)	0.0117	(-)	0.0197
σ_a	-	-0.0802***	-	-0.0810***	-0.0721***
	-	(0.0065)	-	(0.0066)	(0.0074)
σ_0^r	-	-	0.4413**	0.5954***	0.5803***
	-	-	(0.0109)	(0.0202)	(0.0217)
σ_r^r	-	-	-	0.0124***	0.1310***
	-	-	-	(0.0047)	(0.0355)
σ_{rz}^r	-	-	-	-	0.1805**
	-	-	-	-	(0.0492)
σ_{ra}^r	-	-	-	-	-0.0079**
	-	-	-	-	(0.0043)

*Skedastic function estimates for various specifications. These correspond to the elasticities in Table 2. Stars denote level of significance. * $\Rightarrow$ Significant at 10% level, ** $\Rightarrow$ Significant at 5% level. *** $\Rightarrow$ Significant at 1% level.

implications for the optimal R&D policy function. In particular, less productive firms may be more willing to take a chance on doing R&D at the extensive margin. As well, there are option type effects on R&D . Doing more R&D increases the volatility of TFP shocks. Given that firms have convex profits, this increases the average value of the firm next period. Firms will weigh the costs and benefits of R&D taking into account these volatility effects.

My model also clarifies the literature on knowledge capital in important ways. Consider a stock of knowledge capital, c. This literature assumes it accumulates deterministically as:

$$c' = (1 - \delta)c + r$$

Using the notation of my paper, it enters the production function as follows:

$$y = \beta_k k + \beta_n n + \beta_c c + \hat{z}$$

In essence, I am identifying knowledge capital plus actual productivity as "TFP" in this paper:

$$z = \beta_c c + \hat{z}$$

Recognizing that the stock of knowledge capital can be identified as a function of the state variables, this implies that:

$$z' = F(\hat{z}, c(\hat{z}_{-1}, \hat{z}_{-2}, \dots)) + \gamma_r r$$

The existence of non-linearities and the inclusion of the skedastic function reject two aspects of the knowledge capital model:

1. Accumulation of knowledge capital is not linear.

2. Accumulation of knowledge capital is not deterministic.

These findings, along with the difficulties inherent in measuring the stock of knowledge capital support my approach.

To be clear, I am not suggesting that a model of knowledge capital is not a useful way to view R&D . Quite the opposite. I am saying that the underlying process is more complex than this literature generally allows and that this methodology generalizes the concept. Griliches et al. (1979) introduced the idea of thinking of knowledge as a type of intangible capital. This implies that we should think of the dynamic R&D decisions akin to investment. This is exactly what I do in this thesis, taking into account non-linearities and endogenous uncertainty. Firms absolutely make use of R&D investments to try and increase future profits. In the modern economy, intangible forms of capital are increasing in importance.

The stochastic process that I will use in my quantitative analysis is:

$$z' = \mathbb{I}(r = 0)\rho z + \mathbb{I}(r > 0)\rho^r z + \sigma(z, r, a)\xi'$$

$$\sigma(z, r, a) = \sqrt{\sigma_0^2 + \sigma^2 a + \mathbb{I}(r > 0)[\sigma_r^2 r + \sigma_{r,z}^2 rz + \sigma_{r,a}^2 ra]}$$

4.3 Productivity Over the Life Cycle

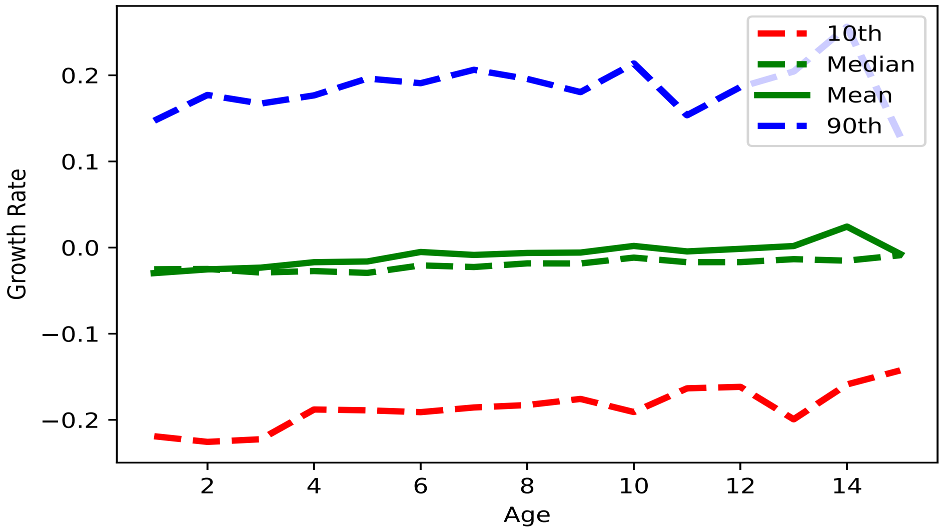
In Section 3.2 I used the Compustat data to document some salient features of the firm life cycle. In particular, I did two exercises to compare to the existing literature on firm dynamics. I constructed the weighted percentiles of the employment growth distribution over the life cycle and I constructed the covariance matrix of log employment. Often these two pieces of evidence are interpreted as evidence of up or out productivity dynamics. However, employment is not the same as productivity. Given the stability of my production function estimates, I can recover the residual of TFP and re-construct these figures.

Figure 4 reconstructs the growth rate distribution over the life cycle. I find that productivity growth is close to flat and slightly increasing. I do not replicate the feature that the variance and skewness of productivity growth is declining over the life cycle.

Next, Figure 5 shows that there is no discernible trend in the standard deviation of TFP over the life cycle. While the employment distribution is fanning out over the life cycle, the TFP distribution displays no such trend. This could be evidence that TFP precedes hiring. Adjustment costs relating to hiring or financing constraints could cause firms with large residual TFP to delay hiring.

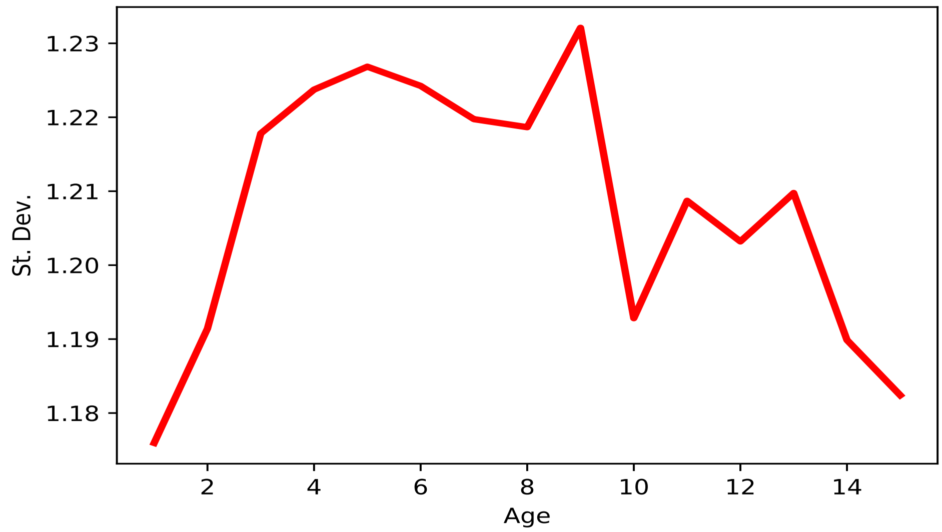
I do find strong evidence of the effects found in Pugsley et al. (2018). The correlation between TFP at age a and $a+j$ does not converge to 0 as j increases. This is suggestive of fixed firm types. It should be noted that I estimated a very similar stochastic process when I included firm fixed effects as opposed to industry dummies. One way to interpret my process for TFP is that it represents the dynamics of productivity after projecting off the firm (or industry) fixed effect and the time dummies (which control for business

Figure 4: The Life Cycle of Productivity Growth



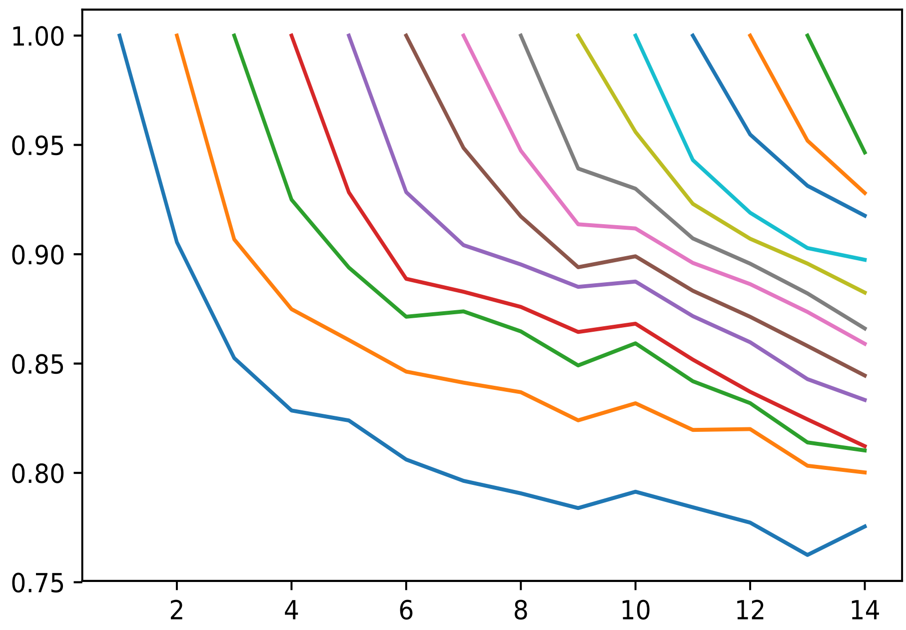
*Using the estimates from my preferred specification I recover productivity as $\exp(z_{i,j,t} + \varepsilon_{i,j,t})$. I then create the weighted percentiles of the TFP growth distribution. I use the average growth measure and average employment as weights.

Figure 5: The Standard Deviation of Productivity Over the Life Cycle



*Using estimates from my preferred specification I recover productivity as $\exp(z_{i,j,t} + \varepsilon_{i,j,t})$. I then compute the standard deviation of productivity over the life cycle.

Figure 6: The Correlation of Productivity Over the Life Cycle



*Using estimates from my preferred specification I recover productivity as $\exp(z_{i,j,t} + \varepsilon_{i,j,t})$. I then compute the correlation between age a and age $a + j$ over the life cycle.

cycle fluctuations). Therefore, using my TFP process generates a specific type of ex-post heterogeneity in productivity outcomes. **Including firm fixed effects complicates the analysis of selection. If the firms know their type right away then they will select very quickly based on this fixed effect. This leads to strong declines in exit rates over the first couple years of firm life. If they learn their type over time this effect is less dramatic. This mechansim requires learning as in Jovanovic (1982).**

5 Research Spending and Market Failure

5.1 Creating the measures of closeness.

TO BE DONE.

- Discussion of the importance of R&D spillovers and literature following Griliches et al. (1979), Hall et al. (2001), Jaffe et al. (1986) and Bloom et al. (2013).
- How do I measure aggregate R&D spillovers? I construct a closeness weighted measure of R&D using patent data. For firm i, the relevant aggregate R&D measure

is:

$$\mathcal{R}_{i,t} = \sum_{j \neq i} \omega_{i,j} r_{j,t}$$

- How do I measure the weights, $\omega_{i,j}$? I follow Bloom et al. (2013), who use patent data. Firms cite over 400 different patent classes when they file a patent. I can view the share of a firms patents which cite field, f. If the vector of these shares is "similar" for firms i and j I interpret them as "close" in innovation space. I will normalize the weights to sum to 1 so that these can be interpreted as weighted averages of log R&D .

- Then I include this measure on the RHS of my TFP processes and analyse the effects.

- I am different from the literature once again because I use flow data on R&D and I examine age interactions. Also, I can try including it in the skedastic function, though these wouldn't be called spillovers in the sense in which the literature considers them.

- My method identifies spillovers in a way that can be included in my model. The fact that these spillovers are outside of the pricing system for R&D leads to a role for public policy. I can use this model to consider:
    1. Across the board subsidies to R&D . These are similar to the R&D and experimentation tax credit which has been used at various times in the US.
    2. Age and size based subsidies to R&D . These policies do not currently exist, to my knowledge. To the extent which young firms innovate "differently" from mature firms, there could be justification for such policies. I may be able to show that conditioning subsidies on age may increase the effectiveness of a given dollar of R&D subsidy. Regardless of what I find, my analysis is disciplined by the empirical methodology in this first chapter.
    3. Age and size based differences in the tax code already exist and may indirectly be influencing innovation policy. I could look at age and size based subsidies which are not related directly to R&D . Given that a lot of R&D is not explicitly expensed, there may be some merits in these indirect subsidies.

5.2 Spillover Results

TO BE DONE.

6 Robustness

- Uncertainty in Parameters of production function: **My strategy for assessing the robustness of my TFP process estimates is as follows. I impose**

constant returns to scale and estimate the different models over a grid of parameter values for $\beta_k = 1 - \beta_n$. For each of these parametrizations I estimate a 95% bootstrap confidence interval and show this in the Box and Whisker Plot Figure. This suggests a range of the relevant parameter space for calibrating my full model which is robust to the validity of my materials inversion.

- Controlling for exit or not? **I did not expect my production function estimates to change when I included exit probabilities. I want to understantd why this is the case.**

- Lag structure of R&D . **What do I find by adding R&D from more than one previous period to my TFP process? Is there an incentive to "keep R&D going"?**

- Materials measure. **Do I find similar estimates using the Compustat measure of wages and salaries? This will be a restricted sample but the data may be "cleaner". Could also introduce a new form of sample bias though, so not clear that it is a better measure.**

- OLS results. **I have included estimates in the appendix for the stochastic process and the skedastic function for the OLS estimates.**

- Spillover weighting matrices. **I need to examine the Jaffe et al. (1986) and Bloom et al. (2013) papers. They list some alternative weighting measures for aggregate R&D spillovers.**

7 Conclusion

This paper was primarily empirical and allowed me to document some interesting facts regarding the uncertainty and heterogeneity of R&D . In particular, I document that firms who do more R&D have more uncertain outcomes and that this uncertainty is heterogeneous.

The next step in this project is to seriously consider the implications of this finding for a model of firm dynamics with selection. That is the intention of the next chapter.

References

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

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*.

Akcigit, U. and Ates, S. T. (2019). What happened to us business dynamism? Technical report, National Bureau of Economic Research.

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

Arrow, K. (1962). The economic consequences of learning by doing. *Review of economic studies*, 29(3):155–173.

Bartelsman, E. J. and Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic literature*, 38(3):569–594.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393.

Buettner, J. (2004). R&d thesis. *LSE Thesis*.

Castro, R., Clementi, G. L., and Lee, Y. (2015). Cross sectoral variation in the volatility of plant level idiosyncratic shocks. *The Journal of Industrial Economics*, 63(1):1–29.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Doraszelski, U. and Jaumandreu, J. (2013). R&d and productivity: Estimating endogenous productivity. *Review of Economic Studies*, 80(4):1338–1383.

Foster, L., Haltiwanger, J. C., and Krizan, C. J. (2001). Aggregate productivity growth: Lessons from microeconomic evidence. In *New developments in productivity analysis*, pages 303–372. University of Chicago Press.

Gandhi, A., Navarro, S., and Rivers, D. (2018). On the identification of gross output production functions.

Griliches, Z. et al. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of economics*, 10(1):92–116.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H., Mairesse, J., and Mohnen, P. (2010). Measuring the returns to r&d. In *Handbook of the Economics of Innovation*, volume 2, pages 1033–1082. Elsevier.

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

Hurst, E. and Pugsley, B. W. (2011). What do small businesses do? Technical report, National Bureau of Economic Research.

Jaffe, A. B. et al. (1986). Technological opportunity and spillovers of r&d: Evidence from firms' patents, profits, and market value. *American Economic Review*, 76(5):984–1001.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, pages 649–670.

Karahan, F., Pugsley, B., and Şahin, A. (2019). Demographic origins of the startup deficit. Technical report, National Bureau of Economic Research.

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

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica: Journal of the Econometric Society*, pages 1263–1297.

Pugsley, B. W., Sedlacek, P., and Sterk, V. (2018). The nature of firm growth.

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

Syverson, C. (2011). What determines productivity? *Journal of Economic literature*, 49(2):326–65.

8 Appendix 1: Data Construction and Sample Selection

The economic variables that are required to estimate a production function are:

- Value Added. $y_{i,j,t}$ in this paper.
  - Value added is theoretically considered to be gross output minus materials.
  - I estimate using a combination of Compustat and SSA data on average wages.
  - I observe SALE which is sales net of discounts and returns. This is a measure of gross revenue.
  - **Net Sales - Material Input** is value added. I need to operationalize these concepts appropriately using the compustat data.
  - I observe Operating Income Before Depreciation (OIBDP) which is SALE - COGS - XSGA. This measure subtracts the employee expense, advertising and marketing expenses and R&D expenses. There is a line item for employee expenses but it is sparsely populated. Therefore, I use EMP x SSA Average Wage as my measure of labour expense. This assumption gives me a bigger, more representative sample of firms. However, it may introduce noise into the measurement of materials and value added if the measurement error of labour expense is not iid.
  - The following relationship holds:

    $$SALE - OIBDP + EMP.EXPENSE + XRD = SALE - MATERIALS = VA$$

    $$MATERIALS = OIBDP - EMP.EXPENSE - XRD$$

    Where any expenses not related to capital accumulation, R&D and paying employees are included in "materials".
  - I deflate value added with the GDP deflator from the BEA. **My understanding is the NBER has 4 digit SIC output deflators. I should be using these in future iterations of this paper.**
- Material Input. $m_{i,j,t}$ in this paper. Discussed in the context of value added above.
- Capital Stock. $k_{i,j,t}$ in this paper.
  - The capital stock is measured as PPEGT which is gross property, plant and equipment. This is measured at cost to the company in the year the capital was purchased.
  - The PPENT is property, plant and equipment net of depreciation. Therefore, PPEGT - PPENT is accumulated depreciation.
  - I have a measure of current depreciation, DP. If I assume constant depreciation then average age of capital is:

    $$\text{AGE}_K = \frac{PPEGT - PPENT}{DP}$$

- I deflate capital stock using the BEA investment price index for the average year that matches the current year minus the average age of capital.

- Employment. $n_{i,j,t}$ in this paper.
  - Employment is easy as the variable EMP in Compustat measures employees.
- Investment. $i_{i,j,t}$ in this paper.
  - I observe CAPX in Compustat. I deflate by the BEA private fixed investment (non-residential) price deflator.
- Research and Development Expenditures. $r_{i,j,t}$ in this paper.
  - XRD measures nominal R&D expenditures in Compustat.
  - Should I deflate with the output or the investment deflator?
- Firm age. $a_{i,j,t}$ in this paper.
  - Firm age is age since listing in this paper.
  - I call the BIRTHYEAR the first time a firm shows up in the data.
  - I call the DEATHYEAR the last time the firm appears in the data.
  - Age is constructed as FYEAR - BIRTHYEAR.
- Dummy variables. ι_i and τ_t in this paper.
  - I create a series of SIC industry dummies at a high level.
  - I create annual year dummies to control for aggregate fluctuations.
  - The value added that I use in the estimation is the residual of the following regression:

    $$y_{i,j,t} = \iota_j + \tau_t + \hat{y}_{i,j,t}$$

    I also get similar results when I run a fixed effects regression, dropping the industry dummy due to collinearity:

    $$y_{i,j,t} = \theta_i + \tau_t + \hat{y}_{i,j,t}$$

    Where θ_i is the firm fixed effect.

I will now enumerate my data dropping procedure for the sample.

1. Start with all observations which have a GVKEY and a FYEAR (a firm id and a year) **541,562**
2. Keep only US firms. **446,223**
3. After the age variable is created, drop observations which are missing sales or employment observations. **338,812**

4. Keep only firms which have fiscal year end in December. **214,818**
5. Drop financial firms, utilities and government entities. **155,530**
6. Drop firms which have negative employment or sales. **149,777**
7. Drop firms with abnormally old capital stocks. Average vintage $<$ 1929. **149,393**
8. Drop observations where there is M&A activity which affects sales by more than 10%. **139,771**
9. Drop before 1962. **134,398**
10. Drop extreme outliers. That is, only keep observations where valueadded, employment, capital and investment is between the 1st and 99th percentiles. There are some extreme outliers which are problematic for interpreting the data. **120,825**
11. Drop observations where OIBDP is missing as we cannot compute value added. **107,365**
12. Drop observations where the lag of either valueadded, capital, employment or investment is missing. This is important due to our estimation strategy. **87,069.**

The sample consists of 83,680 observations on 8,771 firms.

9 Appendix 2: A Model of R&D Spillovers

This section builds on Jaffe et al. (1986), Hall et al. (2001) and Bloom et al. (2013).

Consider an economy with J firms. Each firm $i \in (1, J)$ has a fixed number of innovators, n_i which is a monotonic function of the R&D effort. These scientists are allocated across $\tau \in (1, K)$ fields or technology classes. The allocation is exogenous and represents the "technological profile" of the firm. The total number of scientists is then:

$$n = \sum_{i=1}^{J} \sum_{\tau=1}^{K} n_{i,\tau}$$

There are assumed knowledge transfer probabilities $\omega_{\tau,q}$ for two scientists in fields τ and q.

1. Knowledge transfer only within a field. No cross-field spillovers.
2. Probability of transfer is anonymous to the scientists involved.
3. The probability of transfer is the same for each field.

In sum:

$$\omega_{\tau,q} = \begin{cases} \omega & \text{if: } \tau = q \\ 0 & \text{otherwise} \end{cases}$$

There is an expected number of encounters between scientists of firms i and j which are $n_{i,\tau} n_{j,\tau}$. This implies that:

$$\text{SPILL}_{i,j} = \omega \sum_{\tau=1}^{T} (\frac{n_{i,\tau}}{n_i} \frac{n_{j,\tau}}{n_j}) n_i n_j$$

The aggregate spillovers for firm i are then:

$$\text{SPILL}_{i} = \omega \sum_{j \neq i} \sum_{\tau=1}^{T} (\frac{n_{i,\tau}}{n_i} \frac{n_{j,\tau}}{n_j}) n_i n_j$$

The USPTO has 426 different patent classes. We use this to operationalize this model of spillovers. Firm i has a vector of shares in patent classes:

$$T_i = (T_{i,1}, T_{i,2}, \ldots, T_{i,426})$$

These are measurable. I construct a closeness measure between firms i and j as:

$$\text{TECH}_{i,j} = \frac{P_i P_j^T}{\sqrt{P_i P_i^T} \sqrt{P_j P_j^T}}$$

The spillover measure is then:

$$\text{SPILL}_{i,t} = \sum_{j \neq i} \text{TECH}_{i,j} r_{j,t}$$

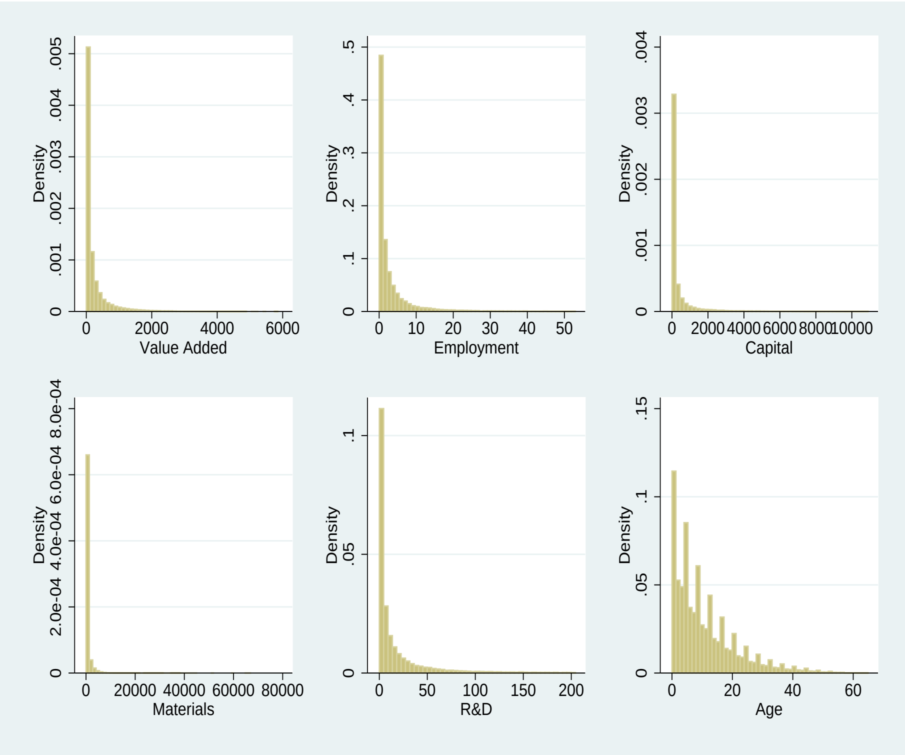
For comparison with my model I call this aggregate R&D and label it as $R_{i,t}$. Therefore, each firm has a closeness weighted measure of aggregate R&D that can be included in the TFP process:

$$z' = g(z, r, R) + \sigma(z, r, R)\xi'$$

For the purposes of my quantitative model I will abstract from these patent classes and assume that the network of R&D spillovers is symmetric. That is, I consider a situation where $R_{i,t} = R_t$ for all firms. In a stationary equilibrium I call it $\mathcal{R}(\mu)$. This is because it depends on the stationary distribution of firms, but is the same for all firms.

10 Appendix 3: Extra Figures and Tables

Figure 7: Histogram's of Key Variables



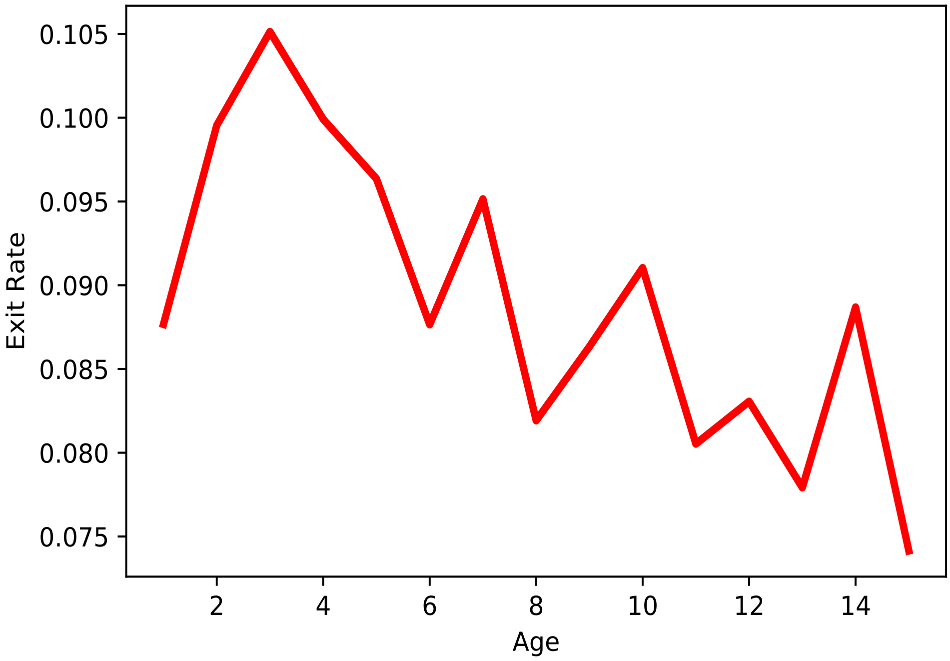
*Value added, capital, R&D , materials and investment are measured in millions of constant 2012 US dollars. Employment is meaasured in thousands of employees and age is measured in years.

Table 5: Production Function Estimates: No Exit Controls

	OLS	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
β_k	0.2571	0.2840	0.2850	0.2763	0.2647	0.2602
	(0.0083)	(0.0195)	(0.0196)	(0.0224)	(0.0190)	(0.0198)
β_n	0.6855	0.6813	0.6758	0.6922	0.7002	0.7013
	(0.0106)	(0.0249)	(0.0242)	(0.0281)	(0.0223)	(0.0234)

*Production function estimates for various specifications without controlling for exit. These correspond to the elasticities in Table 2. Stars denote level of significance. * $\Rightarrow$ Significant at 10% level, ** $\Rightarrow$ Significant at 5% level. *** $\Rightarrow$ Significant at 1% level.

Figure 8: Exit Rate Estimated From Compustat Data



*Exits include being bought out and merged with another company or being dropped from sample due to measurement error. This is why they are "noisy" compared with BDS data.

Table 6: Stochastic Process Estimates: No Exit Controls

	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
γ_0	-0.1224^{***}	-0.1263^{***}	-0.1018^{***}	-0.0933^{***}	-0.0927^{***}
	(0.0161)	(0.0165)	(0.0187)	(0.0167)	(0.0173)
γ_a	-	0.0008^{***}	-	-	0.0007^{***}
	-	(0.0002)	-	-	(0.0002)
γ_{aa}	-	-0.0000^{***}	-	-	-0.0000^{***}
	-	(0.0000)	-	-	(0.0000)
ρ	0.9030^{***}	0.9049^{***}	0.9152^{***}	0.9219^{***}	0.9219^{***}
	(0.0100)	(0.0100)	(0.0127)	(0.0119)	(0.0123)
γ_0^r	-	-	-0.1834^{***}	-0.1749^{***}	-0.1643^{***}
	-	-	(0.0208)	(0.0211)	(0.0226)
ρ^r	-	-	0.8469^{***}	0.8491^{***}	0.8597^{***}
	-	-	(0.0167)	(0.0163)	(0.0181)
γ_r^r	-	-	-	-0.0003	-0.0122^{**}
	-	-	-	(0.0006)	(0.0068)
γ_{rz}^r	-	-	-	-	-0.0114^{*}
	-	-	-	-	(0.0070)
γ_{ra}^r	-	-	-	-	0.0001^{**}
	-	-	-	-	(0.0000)

[*] Stochastic process estimates for various specifications without controlling for exit. These correspond to the elasticities in Table 2. Stars denote level of significance. * ⇒ Significant at 10% level, ** ⇒ Significant at 5% level. *** ⇒ Significant at 1% level.

Table 7: Stochastic Process Estimates: OLS and No Exit Controls

	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
γ_a	-	0.0002^{**}	-	-	-0.0008^{**}
	-	(0.0001)	-	-	(0.0001)
γ_{aa}	-	0.0000^{**}	-	-	0.0000^{**}
	-	(0.0000)	-	-	(0.0000)
ρ	0.9124^{***}	0.9191^{***}	0.9160^{***}	0.9160^{***}	0.9191^{***}
	(0.0091)	(0.0090)	(0.0102)	(0.0102)	(0.0102)
γ_0^r	-	-	-0.0244	-0.0262	-0.0196
	-	-	0.0015	0.0018	0.0019
ρ^r	-	-	0.8704	0.8701	0.8769
	-	-	0.0155	0.0155	0.0164
γ_r^r	-	-	-	0.0013	-0.0010
	-	-	-	0.0005	0.0007
γ_{rz}^r	-	-	-	-	-0.0139
	-	-	-	-	0.0070
γ_{ra}^r	-	-	-	-	0.0001
	-	-	-	-	0.0000

*Stochastic process estimates for various specifications without controlling for exit and estimating production function by OLS. These correspond to the elasticities in Table 2. Stars denote level of significance. * ⇒ Significant at 10% level, ** ⇒ Significant at 5% level. *** ⇒ Significant at 1% level.

Table 8: Skedastic Function Estimates: No Exit Controls

	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
σ_0^2	0.4896^{***}	0.5807^{***}	0.5253^{***}	0.6128^{***}	0.5998^{***}
	(0.0082)	(0.0125)	(0.0113)	(0.0144)	(0.0154)
σ_a^2	-	-0.0076^{***}	-	-0.0077^{***}	-0.0066^{***}
	-	(0.0006)	-	(0.0006)	(0.0008)
σ_0^2	-	-	0.4507^{***}	0.5293^{***}	0.5183^{***}
	-	-	(0.0105)	(0.0157)	(0.0168)
σ_r^2	-	-	-	0.0119^{**}	0.1781^{***}
	-	-	-	(0.0048)	(0.0415)
σ_{rz}^2	-	-	-	-	0.1541^{***}
	-	-	-	-	(0.0424)
σ_{ra}^2	-	-	-	-	-0.0009^{***}
	-	-	-	-	(0.0003)

*Skedastic function estimates for various specifications without controlling for exit. These correspond to the elasticities in Table 2. Stars denote level of significance. * ⇒ Significant at 10% level, ** ⇒ Significant at 5% level. *** ⇒ Significant at 1% level.

Table 9: Skedastic Function Estimates: OLS and No Exit Controls

	AR(1)	Life Cycle	Ext. Margin	Int. Margin	Full Interactions
σ_0^2	0.4900***	0.5798***	0.5279***	0.6137***	0.6010***
	(0.0085)	(0.0125)	(0.0114)	(0.0142)	(0.0151)
σ_a^2	-	-0.0076***	-	-0.0076***	-0.0065
	-	(0.0006)	-	(0.0006)	(0.0008)
$\sigma_{0,r}^2$	-	-	0.4516***	0.5289***	0.5163***
	-	-	(0.0112)	(0.0164)	(0.0174)
σ_r^2	-	-	-	0.0118***	0.0302***
	-	-	-	(0.0049)	(0.0074)
σ_{rz}^2	-	-	-	-	0.1107
	-	-	-	-	(0.0434)
σ_{ra}^r	-	-	-	-	-0.0011
	-	-	-	-	(0.0003)

*Skedastic function estimates for various specifications without controlling for exit and OLS. These correspond to the elasticities in Table 2. Stars denote level of significance. * $\Rightarrow$ Significant at 10% level, ** $\Rightarrow$ Significant at 5% level. *** $\Rightarrow$ Significant at 1% level.