The Dynamics of Firm Profitability, Growth, and Exit

Thomas Buettner

Submitted for the degree of PhD

London School of Economics

July 2004

R 532
10

Abstract

This thesis analyses the dynamics of firm profitability, growth, and exit across different industries.

Chapter 1 documents a striking empirical regularity in the joint distribution of firm profitability and firm size which varies systematically across industries: In industries with a high intensity of R&D investments, there is a strong, systematic "negative tail" of small loss-making firms in the profits-size distribution, whereas this "negative tail" is much less pronounced in industries with low R&D intensity. The chapter also proposes a simple reduced form dynamic model which explains the main empirical features by combining two key mechanisms: a real option effect at the business level and a diversification effect at the firm level.

The second part of the thesis takes a structural approach. Its focus is on estimating the dynamic evolution of firm productivity which is an unobserved state variable in an underlying structural model. In this model, firms make exit decisions and investment decisions in physical capital and in R&D. Chapter 2 extends the model in Olley & Pakes (1996) to include R&D decisions that stochastically affect future productivity realisations and proves that their invertibility approach still applies. It estimates the distribution of future productivity conditional on current productivity and R&D investments, which is the key stochastic primitive in theoretical models of firm dynamics.

Chapter 3 introduces knowledge capital as a second unobserved state variable into the model and extends the invertibility idea and the estimation strategy to the case of two unobserved state variables. Knowledge allows for lagged effects of R&D on productivity while simultaneously accounting for the stochastic nature of R&D. This reconciles the knowledge capital view in the tradition of Griliches (e.g. 1998) with the stochastic approach in the recent literature on firm dynamics.

Contents

List of Tables

List of Figures

Acknowledgements

First of all, I would like to thank my supervisors, John Sutton and Peter Davis for their continuous support, encouragement, patience and kindness throughout the course of my thesis. In their different ways, both have had an immeasurable positive influence on how the final product looks like and I could not have done it without them. John Sutton has tremendously influenced and sharpened my view of what economics can and should do and has taught me the importance of explaining simple empirical regularities on the basis of simple principles. Peter Davis has introduced me to the world of structural modelling with its advantages and disadvantages and has been extremely accessible at almost any hour of the (research) day. Both have provided enormous practical support and the thesis very much reflects their complementary approaches.

I am also very grateful to Mark Schankerman whose comments on my work have always been very valuable and who has encouraged me to keep going forward when I needed it most; to John van Reenen with whom I've had many useful discussions about Chapter 2; to Martin Pesendorfer for challenging the reduced form approach of Chapter 1 and discussions about the second part of the thesis; and to Ariel Pakes with whom I had a brief but very helpful discussion about how to introduce a second control into his model.

Many fellow PhD students and friends have made my time at LSE more enjoyable and intellectually more fruitful, most notably Ralph Bayer, Nick Bloom, Antonio Butta, Jad Chaaban, Barbara Luppi, Maria-Luisa Mancusi, Felix Muennich, Katharina Sailer, Alberto Salvo-Farre, Julia Shvets and many others.

Financial support from the German Academic Exchange Service (DAAD), the Economic and Social Research Council (ESRC), and the Suntory and Toyota Centre for Economics and Related Disciplines (STICERD) is gratefully acknowledged.

Finally, the biggest burden over my thesis years was probably borne by my partner, Laure-Agnès Galtier, by my family, and by my friends inside and outside Academia. They have lived – to varying degrees of intensity – through all the ups and downs with me in this difficult period of my life and their encouragement and emotional support was invaluable.

Introduction

This thesis analyses different aspects of the dynamics of firm profitability, growth, and exit across different industries. It has two parts. The first part (Chapter 1) analyses a striking empirical regularity in the joint distribution of firm profitability and firm size and proposes a simple reduced form dynamic model that explains this feature by combining a real option effect with a diversification effect. The second part (Chapters 2 and 3) takes a structural approach and focuses on estimating the central primitive that drives idiosyncratic outcome paths across firms in theoretical models of firm dynamics: the stochastic evolution of firms' productivity state conditional on current productivity and research and development (R&D). Chapter 2 does this by extending the techniques by Olley & Pakes (1996) to a structural model in which firms invest in physical capital and R&D. Chapter 3 further extends this approach by introducing an additional unobserved knowledge state into the model, which allows to capture lags in the R&D process.

The starting point of the thesis is a striking shape in the joint distribution of firm profitability and firm size: There is a strong "negative tail" in the profits-size distribution of small, highly unprofitable firms, many of which are reporting losses in the same order of magnitude as their asset values. Moreover, this tail disappears for bigger firms. Analysing and explaining the shape of this distribution is the central theme of Chapter 1.

The first part of the chapter provides a detailed empirical characterisation of the profits-size distribution across industries. As this distribution is an endogenous outcome of the underlying dynamics of firms, the chapter also derives stylized facts on firm dynamics in terms of firm profitability, growth, and exit. The focus here is on how the distribution and the underlying dynamics vary across different industries. Using Standard and Poor's COMPUSTAT database on firms listed on North American stock markets across 42 4-digit SIC industries, the key empirical finding of the chapter is that the shape of the profits-size distribution and the dynamics of firms vary in a systematic way between industries with a high intensity of R&D investments and industry with low R&D intensity:

1. The "negative tail" in the distribution of small, loss-making firms is systematically more pronounced in industries with high R&D intensity than in industries with low R&D intensity.

2. The variance of the change in profit rates from year to year is increasing in R&D intensity.

3. Small, unprofitable firms in high R&D industries have lower exit probabilities, higher probabilities to remain unprofitable, and higher probabilities to become profitable than their counterparts in low R&D industries.

The systematic variation of these features with industry R&D intensity indicates that there are underlying driving mechanisms that are common across industries. The key idea of the chapter is that the observed empirical regularities are due to a combination of two effects: an option value effect and a diversification effect.

The option value effect works as follows. Firms in high R&D industries face a high uncertainty about the future evolution of their profits. This is consistent with the fact that the variance in the change in profit rates is increasing in R&D intensity. Therefore, the option value of staying in the industry for a loss-making firm is higher in high R&D industries than in low R&D industries. Loss-making firms in high R&D industries are hence willing to face higher losses before they optimally decide to exit than their counterparts in low R&D industries. This option value effect generates the negative tail in the profits-size distribution of high R&D industries.

The second effect is a diversification effect. Larger firms tend to be more diversified than small firms implying that the profitability of large firms is an average across more (diversified) businesses than the profitability of small firms. This effect leads to a decline in the cross sectional variance of firm profitability with firm size and, crucially, to the disappearance of the negative tail in the profits size distribution for large firms.

The second part of Chapter 1 models these effects in a simple reduced form model of firm dynamics and shows that this model can generate the main features of the profits-size distribution. In the model, a firm consists of a number of independent businesses which arrive randomly over time. At any given point in time each firm decides whether to continue or abandon each of its constituent businesses. The profit flow from each business follows an exogenous stochastic process which gives rise to a simple real option problem. When a business' current profit flow is negative, the firm has to decide whether to abandon the business taking into account the option value that the business may become highly profitable in the future. This option value of continuation increases and the profit flow at which it is optimal to exit decreases with the variance of the stochastic process of the profit flow. At the firm level, this option value effect is combined with a diversification effect through the aggregation across the firm's constituent businesses.

Based on analytical results for the real option problem for a single business, a simulation approach is used to generate profits-size distributions and intra distribution dynamics. The simulation results show that the model can reproduce the overall shape of the profits-size distribution remarkably well. Moreover, by varying a single parameter – the variance of the underlying stochastic process for the evolution of profit rates of individual businesses – the model can generate the qualitative differences in the empirical distributions between high and low R&D industries. That is, the notion supported by the second stylized fact that industries characterized by a high R&D intensity are high risk environments compared to low R&D industries is sufficient to generate the qualitative differences in the profits-size distribution in this model. The model fails, however, to reproduce the stylised fact on intra distribution dynamics that small unprofitable firms in high R&D industries are more likely to remain in this state than firms in low R&D industries.

The model of Chapter 1 is, of course, a reduced form model in that it treats the evolution of the profit flow of businesses and the arrival of new businesses as purely exogenous. Apart from the decision to shut down businesses, the model abstracts from all other decision variables such as investments in R&D or physical capital. These are clearly important decision variables of the firm so that the model is restrictive in this respect. However, while the model falls short of a fully structural behavioural model and is hence unable to identify structural parameters, modelling industry dynamics by a simple reduced form model is a powerful short-cut that allows to focus on the effect of potentially important underlying mechanisms. This is particularly useful for cross industry studies where the aim is to focus on the mechanisms that are common across industries rather than on modelling any specific industry in great detail. The fact that the simple model is capable of reproducing the striking cross industry differences of the profits-size distribution indicates that the real option effect at the business level combined with a diversification effect at the firm level are two key mechanisms in explaining the observed empirical regularities.

How much do ch 2 and 3 add to our understanding of ch 1?

The second part of the thesis, consisting of Chapters 2 and 3, takes a complementary structural approach. It focuses on a subset of industries and models the dynamics of firms in these industries in more detail. In contrast to Chapter 1, the models in this part abstract from the diversification of firms. However, they are much richer in the modelling of firms' decisions. In these models, profitability is driven by a productivity state that evolves stochastically over time and which firms can influence by investing in R&D. Firms grow by investing in physical capital, and can decide to exit if their expected net present value becomes negative. As in Chapter 1, the fact that the evolution of future profits is subject to uncertainty (via the stochastic evolution of productivity) gives rise to an option value of remaining active even if current profits are negative.

The models in Chapters 2 and 3 are in the spirit of the recent theoretical literature on firm dynamics (e.g. Hopenhayn (1992), Ericson & Pakes (1995)). In this literature, the success or failure of a firm in an industry and the dynamics of profitability and growth are typically driven by the stochastic evolution of a firm specific productivity state which may or may not be influenced through R&D investments. The stochastic evolution of productivity conditional on R&D is the key primitive that generates idiosyncratic differences in outcome paths across firms in this literature. The evolution of productivity and the effect of R&D on other ways for method?

the future productivity distribution is therefore at the heart of understanding the dynamics of firms in R&D intensive industries and their incentives to invest in physical capital and R&D. Chapters 2 and 3 provide an empirical framework to estimate the distribution of future productivity conditional on R&D.

The question of the effect of R&D on productivity is, of course, also the theme of a much older and huge empirical literature initiated by Zvi Griliches (e.g. Griliches (1998) for a collection of papers in this tradition). This literature is concerned with estimating the average or expected (private or social) returns to (firm or industry level) R&D. To do so, this literature typically includes a knowledge capital stock in the estimation of a production function. This knowledge capital is constructed from observed R&D investments of firms and captures the effect of R&D on productivity. While estimating the average effect of R&D on productivity is often the best one can do, an analysis of the effect of R&D on the entire distribution of future productivity at the firm level clearly provides a more complete picture and makes explicit the stochastic nature of the outcomes of R&D. Investigating this distribution provides information on the stochastic environment in the industry under study and therefore forms the basis for a better understanding of the (private) incentives to invest in R&D.

why do we want this?

The empirical approach in Chapters 2 and 3 to estimate this distribution builds heavily on the invertibility approach developed by Steven Olley and Ariel Pakes (1996). In their seminal paper on the Telecommunications Equipment industry, they propose a novel semiparametric method for controlling for unobserved productivity differences across firms in the estimation of production functions. Their method overcomes well known biases in OLS production function coefficients due to the simultaneity of input choices with unobserved productivity differences across firms and due to selection through exit.

Their approach relies on the invertibility of the investment policy function generated by a structural single agent model for the dynamics of firms in the industry. In their model, firms have two state variables, capital and productivity. Productivity evolves stochastically and follows an exogenous first order Markov process, while capital is accumulated deterministically through the firms' investment decisions. Firms maximise their expected discounted value by deciding on whether to exit the industry and by choosing the level of investment in physical capital.

The key idea of the invertibility approach of Olley & Pakes (1996) is as follows. If the marginal value of capital is increasing in productivity, then the optimal investment in physical capital conditional on the current capital stock will be increasing in the level of the expected future productivity. Moreover, if expected future productivity is increasing in current productivity, i.e. if there is persistence in productivity realisations, then this implies that the investment policy function generated by the model (i.e. investment as a function of the state variables productivity and capital) is increasing in productivity. This in turn implies, that the policy function can be inverted to express productivity as a function of investment and capital. Pakes (1994) proves that this property holds for their model.

This inverted policy function allows Olley & Pakes (1996) to control for the productivity state, which is unobserved by the researcher, in the estimation. As the policy function is generated by the dynamic structural model, its inverse depends in a complicated way on all the primitives of the model. Therefore, Olley & Pakes (1996) treat unobserved productivity as an unknown function of investment and capital. This allows them to control for unobserved firm productivity nonparametrically without having to solve the structural model explicitly.

On the basis of this idea, Olley & Pakes (1996) provide a semiparametric estimation method to estimate production function coefficients (which are a subset of the parameters of the underlying structural model) in the presence of unobserved productivity differences across firms. Their method proceeds in two stages. The first stage yields consistent estimates of coefficients of variable factors of production as well as estimates of the joint effect of productivity and capital. The second stage then estimates the capital coefficient and, as a by-product, also produces estimates of the firm specific unobserved productivity state over time.

Chapter 2 extends the approach by Olley & Pakes (1996) to a model in which firms can invest in R&D to improve the distribution of future productivity realisations. In this model, firms have two state variables, capital and productivity, and make exit decisions and investment decisions in physical capital and in R&D. While the capital stock in the next period is a deterministic function of current capital and physical capital investments, R&D has a stochastic effect on the future evolution of unobserved productivity – i.e. the Markov process for the dynamics of productivity in Olley & Pakes (1996) is partly endogenised in this model. This model forms the basis for analysing the effect of R&D on productivity dynamics. The main contributions of the chapter are as follows:

First, the chapter proves that, under certain restrictions on the model primitives, the policy function for capital investments generated by the extended underlying structural model is still monotonic in productivity, conditional on capital. The first part of the intuition for this result is the same as in Olley & Pakes (1996): provided the marginal value of capital is increasing in productivity, investment is increasing in expected future productivity. To complete the intuition, expected future productivity needs to be increasing in current productivity even allowing for the endogenous choice of R&D. That is, the model needs to generate some persistence in productivity realisations in the presence of R&D. The proof of the monotonicity property relies on results from the literature on monotone comparative statics (e.g. Topkis (1978), Athey (1995)) which identifies supermodularity and first order stochastic dominance as key properties in generating monotonicity results.

As in Olley & Pakes (1996), invertibility of the investment policy function is a powerful result which implies that the unobserved productivity state can be expressed as a function of capital and investment even in the extended model including R&D. The result is crucial for controlling for the unobservable productivity states in empirical work and forms the basis for jointly estimating the production function coefficients and the unobserved productivity states.

report results for original estimates - e.g.?

Second, the invertibility result for this model with R&D suggests an estimation approach for production function coefficients and for the unobserved productivity states along the lines of Olley & Pakes (1996). The first stage of the Olley-Pakes estimation method which estimates the production function coefficients for the variable factors of production follows directly. The second stage of the estimation procedure requires some further analysis. This is because of the need to control for the expectation of productivity conditional on past information and on survival to consistently estimate the coefficients of the quasi-fixed factors. In a model with R&D, this expectation not only depends on past productivity and on the survival

Explain the model from perspective of firm – what happens in productivity? – R and D?

Pattern in R and D over time for a particular firm?

probability, but also on past R&D investments. Ignoring the impact of R&D may therefore lead to inconsistent coefficient estimates for the quasi-fixed factors. The chapter presents two alternatives to the stage two estimation equation of the Olley-Pakes algorithm. The first suggests using R&D data but is problematic if the level of R&D is endogenous in the sense that it is correlated with potential measurement error in the dependent variable. The second approach does not require any additional data and exploits the fact that we can control for the effect of R&D by estimating stage two as a fully nonlinear nonparametric function of past productivity and current capital. This second modification can be shown to be asymptotically equivalent to the second stage estimation equation originally proposed by Olley & Pakes (1996). This is because, in the context of our model, the survival probability is strictly increasing in capital conditional on productivity. As a result, both stages of the original Olley-Pakes procedure can be directly applied to our model with R&D.

While the proposed solutions depend on the setup of the structural model, the problem of controlling for R&D in the second stage of the estimation is more general: Whenever expected future productivity depends on R&D (i.e. R&D has some effect), one needs to think carefully about how to control for R&D to obtain consistent estimates of the quasi-fixed factors. This point applies independently of whether the approach to proxy for productivity relies on a property generated by an underlying structural model as in Olley & Pakes (1996) or not. Even if the rationale for the productivity proxy is not based on a structural dynamic model, such as the intermediate inputs approach recently proposed by Levinsohn & Petrin (2003), one still needs such a model in the presence of R&D to justify the second of the estimation procedure.

In fact, an alternative interpretation of stage two of the estimation procedure presents itself. Since productivity and capital are the state variables of the model, future realisations of output can be expressed as expected output conditional on productivity and capital plus an error term which is uncorrelated with the current state. This gives rise to a multiple index model with two arguments, productivity and capital, because the first argument, unobserved productivity, is an index of the nonparametric part of the first stage estimation and the observed capital state.

Identification of the capital coefficient in this multiple index model requires some index restrictions. In the context of the model in this chapter, identification is achieved by replacing the second argument, capital, with the capital stock next period which is known in the current period and relevant for future periods.

It is interesting to note that this rationale not only holds for future output, but for all future variables (labour, investment etc.). While the thesis does not pursue this any further, additional estimation equations for these variables could be added to stage two of the estimation to improve efficiency. Since these additional equations would all be nonparametric multiple index models of the same indices, no additional assumptions on the model would be required.

The third and main empirical contribution is that the firm and year specific productivity estimates from the production function estimation are used to analyse the distribution of future productivity conditional on current productivity and R&D. This forms a basis for testing whether the first order stochastic dominance assumptions of the theoretical model (which are standard in the literature on firm dynamics) are satisfied and hence whether the model is accepted by the data. By providing a method of examining the effect of R&D on the entire distribution of future productivity and by quantifying this effect, the chapter also hopes to contribute to the empirical literatures on firm dynamics and on R&D and productivity.

The study uses firm level COMPUSTAT data for the four 3-digit SIC industries "Pharmaceuticals (SIC 283)", "Computer Hardware (SIC 357)", "Telecommunications Equipment (SIC 366)", and "Software (SIC 737)". The sample spans the years 1980 to 2001 and is characterised by high levels of R&D spending and by a considerable degree of exit. In fact, the mean and median levels of R&D investments exceed the corresponding levels of investment in physical capital in each of these industries. This suggests that controlling for R&D as well as survival is potentially important. Estimates for the capital coefficient are sensitive to the specification of the stage two estimation equation. In particular, the point estimates for capital from our fully nonlinear specification are lower than those estimated from the original Olley-Pakes equation in all industries except "Software". While these differences are not statistically significant, it indicates that despite the asymptotic equivalence of the approaches, their finite sample performance may market share for replacement and or demand for new prod incr

differ.

Specification tests proposed in Olley & Pakes (1996) lead us to accept the model for "Pharmaceuticals" and "Telecom Equipment" but to reject it for "Computers" and "Software". Testing whether the future productivity distributions conditional on R&D and current productivity satisfy the first order stochastic dominance property of the model leads to the same conclusion. For the industries "Computers" and "Software", this suggests that using investment as a proxy for productivity does not adequately control for unobserved productivity differences across firms and that the structural dynamic model does not adequately capture the dynamic features of these industries. However, the estimation approach and the model seem to work well for "Pharmaceuticals" and "Telecom Equipment".

what is capital ? Starting off should

Further analysis of the future productivity distribution conditional on R&D and current productivity shows that productivity is more volatile in "Telecom Equipment" than in "Pharmaceuticals" as shown by a higher dispersion of productivity increments. The average elasticity of next period's productivity with respect to R&D is estimated to be around .02 for "Pharmaceuticals" and, depending on the specification, between .007 and .04 for "Telecom Equipment". These effects are low but significant and represent estimates of the short run returns to R&D in terms of productivity from one year to the next.

Chapter 3 takes the analysis a step further and explores an extension to the model and the empirical techniques of Chapter 2. In particular, this extension addresses the fact that the assumptions on the R&D process in Chapter 2 do not allow for lagged effects of R&D on productivity realisations that lie more than one period in the future. This is because the model in Chapter 2 makes an extreme assumption as to how R&D accumulates into productivity which is common in the theoretical firm dynamics literature (e.g. Ericson & Pakes (1995)): Conditional on the current productivity state, R&D investments improve the distribution of the next period's productivity state and hence the next period's payoff. The effect of R&D indirectly also transmits into more distant periods through its effect on next period's productivity due to serial correlation in productivity. However, there is no direct effect of current R&D on productivity realisations beyond the next period. This extreme view of how R&D accumulates raises the question of possible time

process or product R and D ?

lags between the R&D investments and the point in time when the outcome of the R&D process becomes pay-off relevant – especially for some industries such as "Pharmaceuticals".

At the other extreme, there is the view that R&D accumulates deterministically into knowledge capital in the tradition of Griliches. In this literature, knowledge capital becomes an input in the production function and there is no inherent difference in the accumulation of knowledge capital and that of physical capital. Both forms of capital accumulate through investments and depreciate in a deterministic way. This approach can deal with lagged effects of R&D on production to the extent that knowledge depreciates slowly. However, the researcher has to specify a depreciation rate for knowledge to construct the knowledge stock from past R&D investments, which is to some extent arbitrary. Furthermore, the deterministic accumulation approach neglects the inherently stochastic nature of the R&D process.

To reconcile these views, Chapter 3 introduces a knowledge state into the model of Chapter 2. Rather than being a direct input into the production function, the role of knowledge is to improve the distribution of future productivity. The firm's knowledge state is the result of an accumulation process containing stochastic as well as deterministic elements. Knowledge in the next period depends deterministically on the current knowledge state, on R&D investments, and on the stochastic realisation of productivity.

This combination of deterministic and stochastic elements in the accumulation process offers the advantage to allow for lags in the effect of R&D of more than one period while at the same time accounting for the stochastic nature of R&D which results in serial correlation in productivity realisations. In this way, the approach can integrate the advantages of the traditional Griliches view of deterministic accumulation of knowledge through R&D with those of the stochastic accumulation approach of the theoretical firm dynamics literature.

A direct implication of the role of stochastic realisations of unobserved productivity in the accumulation process is that knowledge itself becomes an unobserved state variable in the model. This complicates the theoretical and empirical analysis, as it rules out the construction of knowledge capital from past R&D investments. Therefore, estimation of the model requires techniques that do not rely on specifying a depreciation rate for knowledge to construct knowledge capital. When feasible, the lack of dependence on a depreciation rate for knowledge constitutes an additional advantage over the Griliches approach as this depreciation rate typically has to be chosen in an ad hoc manner.

Moreover, the fact that knowledge and productivity are both unobserved state variables in the model raises an interesting methodological issue as to how to control for these states. The invertibility result of Olley & Pakes (1996) and of Chapter 2 applies in the context of a single unobserved state variable. Chapter 3 explores whether this invertibility idea can still be applied when there are two unobserved states in the model. As emphasised by Olley & Pakes (1996), the feasibility of the invertibility approach depends on the precise structure of the underlying model and needs to be examined on a case by case basis. This structure will be even more critical in the presence of two unobserved state variables. Therefore, Chapter 3 can by no means offer a general answer to this question.

However, the chapter proves that under certain additional assumptions, the invertibility approach can still be applied to the model with unobserved knowledge capital. Under these assumptions, the policy functions for investment and R&D can jointly be inverted to yield unobserved productivity as a unique unknown function of the observed capital state and observed investments in R&D and physical capital. This is a powerful theoretical result which provides a basis for controlling for unobserved state variables in empirical applications.

One example of an empirical application is, again, the semiparametric approach by Olley & Pakes (1996). Chapter 3 extend this approach to the model with two unobserved states and R&D by including R&D in the list of productivity proxies. As before, estimation proceeds in two stages. While conditions underlying the invertibility results are sufficient to identify the coefficients of the variable factors of production in the first stage, an additional condition is required for R&D to be a valid proxy in the second stage of the estimation which estimates the coefficients of quasi-fixed factors of production. Unfortunately, this is a condition on the second derivative of the expected future value of the firm, i.e. not a condition on the model primitives but the solution of the dynamic model. Therefore, this condition cannot be checked without specifying all the model primitives and solving the dynamic programme – the difficult task the semiparametric approach aims to avoid in the first place. If one wants to follow the semiparametric approach, one can only hope that this condition is satisfied and proceed with the estimation.

Using the same dataset as for Chapter 2, the production function estimates for the model with unobserved knowledge change somewhat, but not dramatically compared to the estimates of that chapter. The labour coefficients in the model with unobserved knowledge tend to be a little higher than that of the previous chapter, while the estimated capital coefficients tend be a little lower. This indicates that if the true model includes knowledge capital, the OLS biases in the labour coefficient due to simultaneity of the input choices may be quantitatively somewhat less important than suggested by the results of the previous chapter, even though they are still present.

Specification tests reject the model with knowledge capital for all industries except "Pharmaceuticals". In that sense, introducing knowledge in the model and including R&D in the list of proxies does not improve on the model rejections in the second chapter. In fact, for "Telecom Equipment" one of the estimation specification in Chapter 2 was accepted, while both knowledge capital specifications are rejected in Chapter 3. The rejections suggest that for these industries investment and R&D are insufficient to proxy for unobserved differences in productivity and knowledge.

reject general model
but accept special case?

Nevertheless, it is interesting to examine the effect of R&D and knowledge on productivity dynamics. The estimate for the short run elasticity of productivity with respect to R&D is around .016 in Pharmaceuticals and around .03 in "Telecom Equipment". These estimates are quite close to the estimates of Chapter 2. The elasticity of productivity with respect to knowledge is estimated to be around .4 in "Pharmaceuticals", while it is much lower in the other industries. As knowledge capital allows for lagged effects of R&D on productivity, this confirms the prior that there are typically long lags in the Pharmaceutical Industry between the innovation of a new drug and bringing it to the market. It also suggests that introducing unobserved knowledge capital is an important improvement for modelling this industry.

The analysis in the second part of the thesis develops an empirical framework for the estimation of the evolution of productivity conditional on current productivity, R&D, and knowledge capital. This is the key stochastic primitive in the models in Chapters 2 and 3 that drives the dynamics of firm profitability, growth, and exit. The advantage of the semiparametric approach in these chapters is that it is relatively easy to implement and computationally not very intensive. It also does not require functional form assumption for the productivity distribution or for the other primitives of the model. Instead, all it requires are assumption on certain properties of primitives coupled with a functional form assumption for the production function.

However, since the approach falls short of estimating the full dynamic models, one cannot conclude on the basis of the estimates of Chapters 2 and 3 alone whether the dynamic structural models of these chapters can generate the profits-size distributions and the dynamics documented in Chapter 1. This would require the specification of all the primitives of the models and the (simultaneous) estimation of all the parameters. Such an estimation could, for example, be implemented using a nested fixed point approach. While the invertibility results of Chapters 2 and 3 could be directly applied to deal with unobserved state variables such as productivity and knowledge in such a context, a nested fixed point estimation is very complicated and computationally very demanding and is beyond the scope of this thesis.

Chapter 1

Losses Not Losers: The Profits-Size Distribution and a Simple Real Option Model

start pt – investment but not uncertainty

other alternatives?

measurement errors?

can we use share prices to test other explanations?

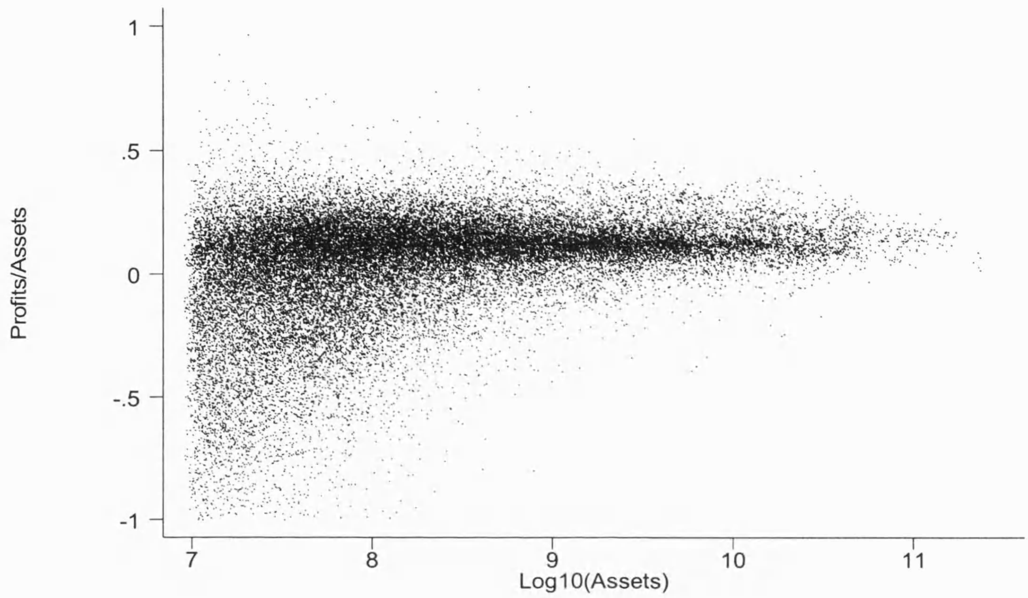
1.1 Introduction

The starting point of this thesis is a striking regularity in the joint distribution of firm profitability and firm size: Figure 1-1 shows a scatter plot of profitability (as measured by operating profits divided by the value of total assets) against firm size (measured by the base 10 logarithm of the value of total assets). Two main features are apparent from this scatter plot. First, the variance of profit rates decreases with firm size. Second, and more surprisingly, there is a long tail of small firms making very high losses – some of them in the same order of magnitude as their asset values – but there are almost no large loss-making firms.

This "negative tail" of small loss-making firms in the profits-size distribution immediately raises a number of questions: Does the shape of the distribution vary across industries? What are the underlying dynamics of firms driving this tail in distribution? How can these dynamics be modelled? This chapter investigates these questions empirically and proposes a simple real option model that can rationalise the main empirical features.

The first part of this chapter provides an empirical characterisation using Stan-

Figure 1-1: The joint distribution of firm profitability and firm size (pooled 1990-2002)



dard and Poor's COMPUSTAT database of the joint distribution of firm profitability and firm size and of the underlying firm dynamics in terms of firm profitability, growth, and exit. The focus here is on how the profits-size distribution and the underlying dynamics differ across industries. The key empirical finding of this chapter is that the shape of the profits-size distribution and the dynamics of firms vary across industries in a systematic way:

1. The "negative tail" in the distribution of small, loss-making firms is systematically more pronounced in industries with a high intensity of investments in R&D than in industries with low R&D intensity.

2. The variance of the change in profit rates from year to year is increasing in the R&D intensity supporting the idea that high R&D industries are "riskier" environments with respect to the future evolution of profits.

3. Small, unprofitable firms in high R&D industries have lower exit probabilities, higher probabilities to remain unprofitable, and higher probabilities to become profitable than their counterparts in low R&D industries.

These stylized facts provide an interesting if crude characterisation of cross industry differences in intra-distribution dynamics and the profits-size distribution.

They also are also a good benchmark for any dynamic model trying to explain these statistical regularities.

Rule out upfront model without uncertainty

The key theoretical idea in this study is that the statistical regularities are due to a combination of two effects: a real option effect and a diversification effect. Compared to low R&D industries, high R&D industries are high risk environments with a high uncertainty about the evolution of future profits. This implies that unprofitable firms in high R&D industries have a high upside risk of becoming very profitable in the future and, therefore, a high option value of remaining active even if they are currently making losses. This effect explains the negative tail in the profits-size distribution and the difference in its length across in low and high R&D industries. The disappearance of the negative tail for larger firms and the decline in the cross sectional variance in profit rates with firm size is driven by a higher degree of diversification of big firms.

The second part of this chapter models these two effects in a simple reduced from dynamic model. In the model, a firm consists of a number of independent businesses which arrive randomly over time. The only decision a firm can make at any given point in time is whether to continue or abandon each of its constituent businesses. The profit flow from each business follows an~~d~~ exogenous stochastic process which gives rise to a simple real option problem. When a business' current profit flow is negative, the firm has to decide whether to abandon the business taking into account the option value that the business may become highly profitable in the future. At the firm level, this option value effect is combined with a diversification effect through the aggregation across the firm's constituent businesses.

Based on analytical results for the real option problem for a single business, a simulation approach is used to generate profits-size distributions and intra distribution dynamics. The simulation results show that the model can reproduce the overall shape of the profits-size distribution remarkably well. Moreover, by varying a single parameter – the variance of the underlying stochastic process for the evolution of profit rates of individual businesses – the model can generate the qualitative differences in the empirical distributions between high and low R&D industries. That is, the notion supported by the second stylized fact above that industries characterized by a high R&D intensity are high risk environments compared to low R&D industries is sufficient to generate the qualitative differences in the profits-size distribution in this model. However, the model fails to reproduce the most striking stylised fact on intra distribution dynamics that small unprofitable firms in high R&D industries are more likely to remain in this state than firms in low R&D industries.

why? should have higher option value!

The model in this chapter is, of course, a reduced form model in that it treats the evolution of the profit flow of businesses and the arrival of new businesses as purely exogenous. Apart from the decision to shut down businesses, the model abstracts from all other decision variables such as investments in R&D or physical capital. These are clearly important decision variables of the firm so that the model is restrictive in this respect.[1] However, while the model falls short of a full structural behavioural model and is hence unable to identify structural parameters, modelling industry dynamics by a simple reduced form model is a powerful short-cut that allows the researcher to focus on the effect of potentially important underlying mechanisms that are common across industries.

The true dynamics within an industry and the resulting profits-size distributions are certainly much richer and differ along more dimensions than the simple model can accommodate. However, the model goes a long way in explaining cross industry differences in the profits-size distributions. This is an indication that the real option effect at the business level combined with a diversification effect at the firm level are two key mechanisms in explaining the observed empirical regularities.

The "negative tail" in the profits-size distribution has, to my knowledge, not been previously recognised or modelled in this form. There is, however, a large body of related literature. On the empirical side, the dynamics of firm profitability and the persistence of profits is the subject of a long literature initiated by Dennis Mueller's seminal book on firms' long run profit rates (Mueller (1986); see also Mueller (1990) for a cross country collection of studies in this tradition). The typical finding in this literature is that there are persistent differences in long run profit rates across firms and that the adjustment process towards these long run

[1] One interpretation would be that decision variables such as R&D have already been optimised out.

rates is relatively quick.

Anita McGahan (1999) takes a different approach in that she analyses the evolution of profit rates of businesses over time. She focuses on transition probabilities between high, medium, and low profitability states over an number of years and examines individual company histories to rationalise the observed transitions. Her study raises some interesting dynamic issues but fails to address what in the light of this chapter are important factors. In particular, her study does not address the question of industry specific factors or of the distribution of profit rates and fails to recognise the importance of the size dimension.

The recent theoretical literature on firm dynamics can be broadly classified in two major strands. The first strand, initiated by Boyan Jovanovic (1982) is a class of learning models in which firms are endowed with a time invariant productivity parameter which is unknown to the firms. Firms learn their own productivity over time through their profit realisations so that low productivity firms will exit over time. This theoretical literature is in the spirit of the empirical studies in the tradition of Mueller (1986) in the sense that firms are endowed with time invariant long run profit rates.

The second strand in the theoretical literature can be described as Markovian models. Here a firm's state variable evolves over time and this evolution only depends on the current state and current decisions. In this class fall the contributions of Hugo Hopenhayn (1992) for a single agent models and the papers by Richard Ericson and Ariel Pakes (1995) and their followers for multiple agent models with interactions and R&D investments. The stochastic model in this chapter is a simple Markovian single agent model in this class and can be interpreted as the reduced form of a richer model that explicitly models investment decisions.

The theoretical techniques employed in this chapter are very much motivated by the pioneering work on real options by Avinash Dixit and Robert Pindyck (see for example Dixit & Pindyck (1994)). While the tradition in the industrial organisation literature is to model firms in discrete time, the model proposed here is a continuous time Brownian motion model. The advantage of this approach is that analytical solutions to the firm's decision problem are available for special cases.

The approach of modelling firms as consisting of (approximately) independent businesses is relatively recent. John Sutton (1998) demonstrates that this leads to very accurate predictions on the minimum degree of skewness in the size distribution of firms. The approach also proves very powerful in explaining the scaling relationship in the decline of the variance of firm growth rates with increases in firm size (Sutton 2002).

The chapter is organised as follows. The first part (section 1.2) documents the empirical regularities. First, the data is introduced. Then cross industry differences in the profits-size distribution are analysed visually and econometrically. The first part concludes with the derivation of a set of stylized facts on intra-distribution dynamics. The second part of the chapter (section 1.3) introduces the model, derives analytical results for the firms' real option problem and uses simulations to compare the models predictions with the stylized facts developed in the first part. Section 1.4 concludes.

1.2 Empirics

1.2.1 Data

This study uses an unbalanced panel of firm level data for 42 4-digit SIC industries in non-financial sectors over the period 1990 - 2002. The dataset is constructed from Standard and Poor's COMPUSTAT database. COMPUSTAT contains accounting data and stock market information on firms listed on North American stock markets that submit reports to the US Securities and Exchange Commission (SEC).[2] Firms with total assets below 10 million US$ submit these reports voluntarily and are deleted from the sample to avoid selection issues.

Table 1.1 lists the industries in our sample and the number of firms and firm years in each industry. As the industry R&D intensity will turn out to be highly correlated with the negative tail in the distribution, industries are classified in two groups on the basis of their R&D intensity defined as the ratio of industry

[2] Detail on the selection of industries and the construction of firm and industry variables are given in Appendix A1.1.

we don't know which firms in the R&D ind actually do R&D

R&D expenditure to industry sales: There is an experimental group of 18 high R&D industries with R&D/Sales ratio exceeding 4%, and a control group of 20 low R&D industries with R&D intensity below 1%. The table also reports the R&D intensity and the advertising to sales ratio for each industry as well as the percentage of non-missing observations on R&D and advertising expenditure.[3]

The main firm variables for this study are the book values of operating income before depreciation, total assets, and total sales. These variables were inflated/deflated to constant 1996 US$.[4] Operating profits consist of the firms' sales revenue net of costs of goods sold including overheads and R&D expenditure. We use three measures of profitability: Operating profits divided by assets ("profits/assets"); operating profits divided by total sales ("profits/sales"); and operating profits gross of R&D expenditure divided by total assets ("gross profits/assets"). The size measures employed are the base ten logarithm of the firm's total assets or sales (in real terms) respectively. Table 1.2 contains summary statistics on these variables for groups of high R&D and low R&D industries for the years 1990 to 2002. Firms in the high R&D subsample are, on average, much smaller in terms of total assets, sales revenue and number of employees than firms in the low R&D subsample. While their mean and median values for profits/assets are lower than for low R&D firms, their average gross profit rates are slightly higher.

Exit is a quantitatively important phenomenon in the data and firms exiting the dataset are assigned a reason for deletion by COMPUSTAT. The main exit reasons by number of occurrences are "merger and acquisition", "bankruptcy", "liquidation", and a category named "other reasons" which includes companies that have stopped reporting to the SEC. This last category is most prominent for

[3]The table also includes a column reporting the R&D intensity from the US Federal Trade Commission's Line of Business data set of the 1970's for manufacturing industries. The R&D data of this data set is of much better quality than the Compustat data, but uses a slightly different industry definition. We only report numbers for which a reasonably close match of the industry definitions could be found. While the data cover different time periods and come from very different sources, the R&D intensity figures are remarkably close. In particular, with regards to our classification in high and low R&D industries, none of the industries for which the alternative figure is available would have been classified differently on the basis of the Line of Business R&D intensity with the exception of "Surgical and Medical Instruments and Apparatus SIC 3842" which would have fallen in the excluded set of medium R&D industries.

[4]While accounting profits are at best a noisy measure of economic profits, they are the only measure of current economic profits available. For a discussion of accounting profits versus economic profits see e.g. Mueller (1990, Appendix to chapter 1)

Table 1.1: Industries, R&D intensity and advertising intensity, 1990-2002.

				R&D intensity			Advert.int.	
a. High R&D industries (R/S > 4%)	**SIC**	**# firms**	**# obs**	**R/S**	**R/S LoB**	**rep. rate**	**A/S**	**rep. rate**
PHARMACEUTICAL PREPARATIONS	2834	286	1989	11.2%	10.2%	93.4%	3.7%	35%
IN VITRO,IN VIVO DIAGNOSTICS	2835	86	571	23.6%		98.7%	0.6%	29%
BIOLOGICAL PDS,EX DIAGNSTICS	2836	179	1100	37.4%		95.1%	0.5%	12%
ELECTRONIC COMPUTERS	3571	57	329	6.6%	8.9%	95.6%	1.4%	51%
COMPUTER STORAGE DEVICES	3572	50	356	8.0%		96.8%	0.6%	47%
COMPUTER COMMUNICATION EQUIP	3576	115	698	18.0%		96.8%	0.4%	46%
TELE & TELEGRAPH APPARATUS	3661	127	828	10.7%	4.9%	94.8%	0.1%	38%
RADIO, TV BROADCAST, COMM EQ	3663	128	906	9.6%	4.9%	96.0%	0.2%	26%
SEMICONDUCTOR,RELATED DEVICE	3674	192	1318	12.5%	6.1%	98.0%	1.6%	27%
ELEC MEAS & TEST INSTRUMENTS	3825	67	480	12.3%	4.8%	94.9%	1.1%	38%
LAB ANALYTICAL INSTRUMENTS	3826	52	334	8.0%	4.8%	98.0%	0.5%	40%
SURGICAL,MED INSTR,APPARATUS	3841	98	578	6.7%	3.8%	97.8%	0.2%	20%
ORTHO,PROSTH,SURG APPL,SUPLY	3842	88	587	4.3%		89.3%	0.7%	32%
ELECTROMEDICAL APPARATUS	3845	138	864	10.0%		99.2%	0.4%	33%
CMP PROGRAMMING,DATA PROCESS	7370	356	1539	6.1%		53.2%	0.9%	36%
PREPACKAGED SOFTWARE	7372	724	3733	15.1%		92.9%	1.7%	42%
CMP INTEGRATED SYS DESIGN	7373	230	1291	6.6%		77.1%	0.2%	29%
COML PHYSICAL, BIOLOGCL RESH	8731	68	323	12.5%		72.6%	0.1%	13%
b. Low R&D industries (R/S < 1%)	**SIC**	**# firms**	**# obs**	**R/S**	**R/S LoB**	**rep. rate**	**A/S**	**rep rate**
GOLD AND SILVER ORES	1040	107	738	0.0%		5.2%	0.0%	3%
CRUDE PETROLEUM & NATURAL GS	1311	329	2195	0.3%		9.9%	0.0%	7%
OPERATIVE BUILDERS	1531	57	491	0.0%		5.7%	0.6%	46%
PETROLEUM REFINING	2911	69	577	0.5%	0.3%	49.2%	0.0%	8%
STEEL WORKS & BLAST FURNACES	3312	70	581	0.5%	0.4%	31.4%	0.0%	5%
TRUCKING, EXCEPT LOCAL	4213	65	517	0.0%		2.3%	0.0%	2%
AIR TRANSPORT, SCHEDULED	4512	61	458	0.0%		1.6%	1.2%	67%
RADIOTELEPHONE COMMUNICATION	4812	126	686	0.5%		17.7%	1.7%	49%
TELEVISION BROADCAST STATION	4833	66	437	0.4%		8.4%	1.9%	40%
ELECTRIC SERVICES	4911	184	1779	0.0%		0.0%	0.0%	0%
NATURAL GAS DISTRIBUTION	4924	59	603	0.0%		0.0%	0.0%	0%
ELECTRIC & OTHER SERV COMB	4931	83	867	0.0%		0.1%	0.0%	0%
COMPUTERS & SOFTWARE-WHSL	5045	61	387	0.0%		69.4%	0.0%	16%
GROCERY STORES	5411	82	621	0.0%		56.3%	0.6%	55%
EATING PLACES	5812	185	1304	0.0%		70.9%	2.8%	68%
CATALOG, MAIL-ORDER HOUSES	5961	98	523	0.5%		62.1%	5.4%	54%
HOTELS AND MOTELS	7011	72	442	0.0%		54.3%	1.1%	54%
HELP SUPPLY SERVICES	7363	65	452	0.1%		14.8%	0.2%	19%
MEDICAL LABORATORIES	8071	58	351	0.7%		34.5%	0.1%	16%
ENGINEERING SERVICES	8711	57	414	0.1%		16.0%	0.0%	3%
c. Medium R&D industries (1% < R/S < 4%)	**SIC**	**# firms**	**# obs**	**R/S**	**R/S LoB**	**rep. rate**	**A/S**	**rep rate**
PERFUME,COSMETIC,TOILET PREP	2844	50	319	1.7%	2.5%	65.8%	7.6%	69%
MOTOR VEHICLE PART,ACCESSORY	3714	92	671	2.2%	1.0%	66.0%	0.1%	19%
PHONE COMM EX RADIOTELEPHONE	4813	277	1793	1.2%		16.7%	0.7%	23%
CABLE AND OTHER PAY TV SVCS	4841	89	482	1.2%		12.2%	1.3%	49%

"R/S" and "A/S" denote the industry R&D (resp. adverting) to sales ratios computed from Compustat.
The column labelled "R/S LoB" reports the corresponding industry R&D intensity from the FTC's Line of Business data.
Columns labelled "rep.rate" give the percentage of nonmissing R&D (resp.) advertising data in Compustat.

Table 1.2: Summary statistics by subsample, 1990 - 2002.

a. High R&D Industries

	# obs.	Mean	Std.D.	Min	Max	Median	Units
Profits	17824	108.84	699.64	-5248.74	17626	2.06	10^6
Assets	17824	767.27	4106.72	9.04	103135	59.10	10^6
Sales	17824	635.33	3451.53	-6.06	83628	43.15	10^6
Employees	16254	3.14	14.95	0.00	373.82	0.30	10^3
R&D	16117	73.94	342.45	-0.15	5680.08	8.59	10^6
Log10(Assets)	17824	7.93	0.71	6.96	11.01	7.77	
Log10(Sales)	17427	7.71	0.90	2.97	10.92	7.66	
Profits/Assets	17824	-0.04	0.33	-5.42	4.78	0.05	
Profits/Sales	17430	-4.68	100.78	-9918.00	26.63	0.06	
Gross Profits/Assets	17430	-1.28	43.05	-5053.00	895.96	0.18	

b. Low R&D Industries

	# obs.	Mean	Std.D.	Min	Max	Median	Units
Profits	14423	358.02	1297.87	-2411.89	31162	39.47	10^6
Assets	14423	2953.82	9343.40	9.23	233258	368.76	10^6
Sales	14423	2175.30	8189.34	-24.53	192801	312.87	10^6
Employees	13091	9.47	34.02	0.00	779.10	1.60	10^3
R&D	3329	23.08	96.30	0.00	1162.07	0.00	10^6
Log10(Assets)	14423	8.62	0.89	6.97	11.37	8.57	
Log10(Sales)	14232	8.47	0.95	3.26	11.29	8.51	
Profits/Assets	14423	0.09	0.19	-12.62	1.01	0.12	
Profits/Sales	14236	-1.26	51.37	-3713.00	3.22	0.14	
Gross Profits/Assets	14236	-1.25	51.35	-3713.00	3.22	0.14	

Table 1.3: Exit rates for firms present in 1990

State by 2002	#	%
Survivors	958	49.69
M&A	707	36.67
Bankr/Liqu	74	3.84
Other Exit.	189	9.8
	1928	100

small companies. For our purposes, exit reasons are grouped into three categories: "merger and acquisition", "bankruptcy and liquidation", and "other". While bankruptcy and liquidation are clearly events indicating the failure of a company, "merger and acquisition" can be a success or a failure. We believe that the reason "other" is also a failure category. Of the 1928 firms that were active in 1990, 37% exit through "merger and acquisition" by 2002, about 4% go bankrupt or liquidate, 10% exit through other reasons and 50% survive (Table 1.3). Table 1.4 gives the number of firms, entrants, and exits by year for an augmented data set back to 1980. A striking feature from this table is that the number of entrants and of merger and acquisitions has been much higher in the 1990s than in the 1980s.[5]

[5]The very low number of entrants and exiting firms in the years post 2000 may be due to lags in the updating procedure of the database for these firms as well as due to a slowdown of IPO activity in this period.

Table 1.4: Entry and exit by year and exit reason

	# Firms	# Entrants	Survivors	M&A	Bankr/Lqn	Other
1980	1,154	1,154	1,106	35	5	8
1981	1,236	140	1,185	23	5	23
1982	1,425	246	1,372	26	5	22
1983	1,573	208	1,476	46	17	34
1984	1,590	124	1,468	55	15	52
1985	1,651	185	1,553	59	10	29
1986	1,731	191	1,621	73	9	28
1987	1,791	183	1,669	75	10	37
1988	1,797	141	1,679	64	14	40
1989	1,798	134	1,715	40	18	25
1990	1,861	154	1,771	45	19	26
1991	1,975	212	1,897	35	15	28
1992	2,163	268	2,086	31	8	38
1993	2,370	286	2,252	78	15	25
1994	2,534	311	2,377	121	9	27
1995	2,870	498	2,719	113	13	25
1996	3,154	436	2,926	174	24	30
1997	3,201	304	2,924	221	25	31
1998	3,362	478	3,059	241	21	41
1999	3,467	454	3,114	255	6	92
2000	3,331	257	3,081	203	2	45
2001	2,933	52	2,811	113	3	6
2002	2,291	19	2,276	14	0	1

1.2.2 The cross sectional profits-size distribution

The shape of the joint distribution of firm profitability and firm size in Figure 1-1 has two striking features: A decline in the variance of firm profitability with firm size; and a long tail of small loss-making firms in the distribution. While the first feature is consistent with a standard diversification argument, the second is much more striking: There is a "negative tail" of small firms making very high losses which disappears relatively quickly with increasing firm size and there are almost no loss making firms with asset values above US$ 1 billion (10^9). This negative tail in the distribution is the central object of this study.

First, we briefly examine changes in the pooled distribution over time. We then move to the main empirical part of the chapter and analyse cross industry differences in the distribution graphically and econometrically using quantile regressions.

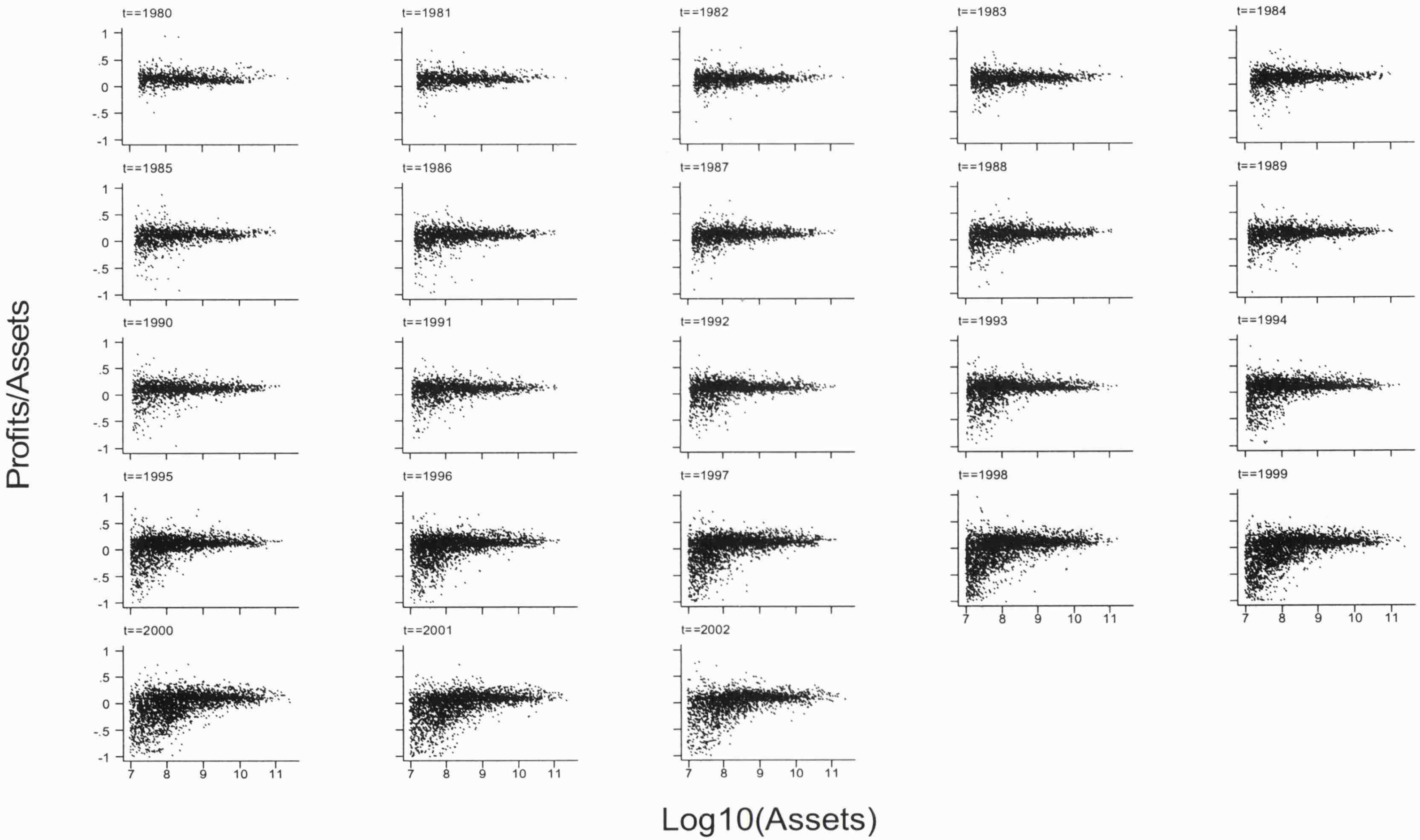
The cross sectional distribution over time

Before investigating differences in the profits-size distribution across industries, it is instructive to analyse, whether the pooled profits-size distribution is stable over time. To do so, we use an augmented data set over the period 1980 to 2002. Figure 1-2 shows scatter plots of the joint distribution of profits/assets and log assets for each year. It emerges that the negative tail in the distribution gets stronger over time. It is (almost) inexistent in the early 1980s, then slowly appears and grows strong in the second half of the 1990s.

There are two candidate explanations for this. The first is that the negative tail is present in the early years in this sample but that it cannot be observed because the real value of the nominal reporting threshold in assets declines over time. The nominal threshold of 10 million US$ has approximately halved over the extended sample period from 17.5 million 1996 US$ (or $10^{7.24}$) in 1980 to only 9.0 million 1996 US$ (or $10^{6.96}$) by 2002. This suggests that the lack of a negative tail in the distribution in the 1980s may be partly due to this inflation induced change in the reporting requirements. However, it is unlikely that this effect fully accounts for the emergence of the negative tail in the 1990s.

A second candidate explanation is that the negative tail may reflect the fact

Figure 1-2: The joint cross-sectional distribution of firm profitability and firm size over time



that access to financial markets has become easier and more popular for small firms over the 1990s. This easing of access to finance for small firms is mirrored by the number of entrants into the dataset over time, where entrants are firms that enter the stock market through IPO's or that for the first time exceed the reporting thresholds. While the average number of entrants in the data from 1981 to 1990 is 182.5 per year, it increases to an average 364.4 per year form 1991 to 1998 (see Table 1.4 for the number of entrants by year). So the fact that access to financial markets has become easier and more popular for small firms over the 1990s may also have played a role in the emergence of the "negative tail".

The evolution of the cross sectional distribution over time constitutes an interesting development. However, given the selection effect operating through the reporting threshold it is difficult to pin down the factors driving this pattern by reference to the present dataset. The focus of this study is instead confined to analysing the profits-size distribution across different industries, rather than over time. For this reason, we will restrict attention to the period after 1990 in the analysis that follows and treat the cross sectional distribution as being stable over this sample period.

The cross sectional distribution by industry

well done

The key empirical finding of this study is that the joint distribution of profitability and size differs across industries in a systematic way. We argue that the negative tail is systematically stronger in industries with high R&D intensity than in low R&D industries. Figure 1-3 shows scatter plots of profits/assets against log assets by industry (pooled over time) for the subsample of industries classified as high R&D industries. A visual inspection of Figure 1-3 reveals that the joint distributions of profits over assets and log assets exhibit a long negative tail in practically all 18 high R&D industries. This negative tail appears particularly strong in the pharmaceutical and biotech industries (SIC's 2834-2836) and in software industries (SIC's 7370-7373).

Figure 1-4 shows the distributions for the set of low R&D industries. In contrast to the high R&D distributions, most of these distributions do not show a clearly discernible negative tail or, at least, the negative tail appears much weaker than for most high R&D industries. The exception seems to be "Catalog, Mail-Order Houses (SIC 5961)".[6] ok

Figures 1-5 and 1-6 show the corresponding scatter plots of the alternative performance and size measures profits/sales and log sales. Although the dispersion of this measure of profitability tends to be higher, the general impression arising from the pictures is the same as before: While most high R&D industry distributions have a clearly distinguishable negative tail (Figure 1-5), only a few of the low R&D industries distributions exhibit this feature (Figure 1-6). There is a large number of highly unprofitable small firms in the low R&D industries "Gold and Silver Ores – SIC 1040" and "Crude Petrol and Natural Gas Extraction – SIC 1311", but this seems to be compensated by a much higher dispersion of profit/sales in these industries than in others.

Finally, Figures 1-7 and 1-8 show the corresponding scatter plots for gross profits/assets. Using profit rates gross of R&D expenditure by the firms clearly shifts up the distribution of profit rates and has a stronger impact on high R&D

[6] "Catalog, Mail-Order Houses (SIC 5961)" includes a number of internet startups such as Amazon.com. However, screening out firms with ".com" in their names did not remove the impression of a strong negative tail in this industry.

Figure 1-3: Profits/assets against log assets by industry - High R&D Industries

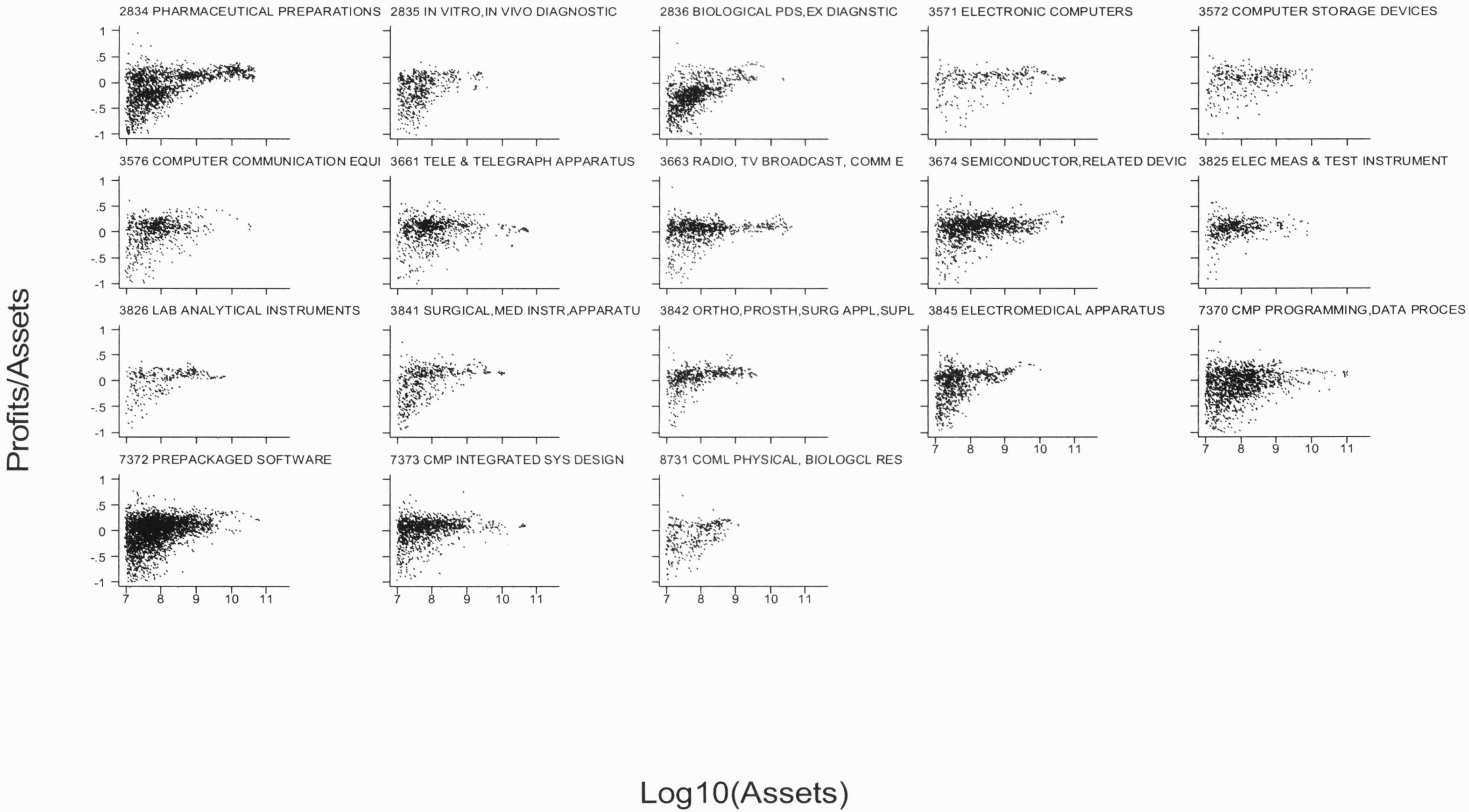


Figure 1-4: Profits/assets against log assets by industry - Low R&D Industries

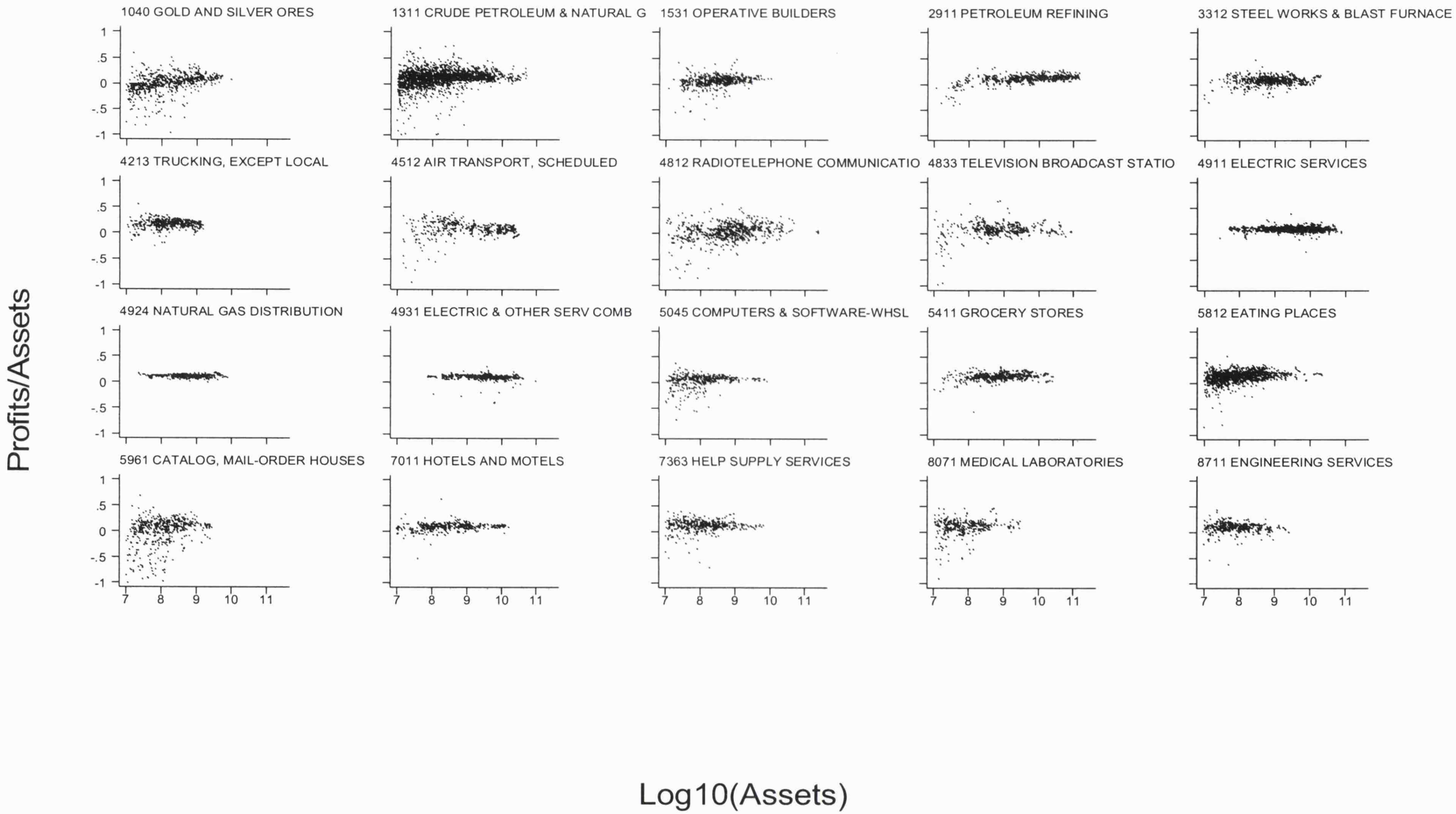


Figure 1-5: Profits/sales against log sales by industry - High R&D Industries

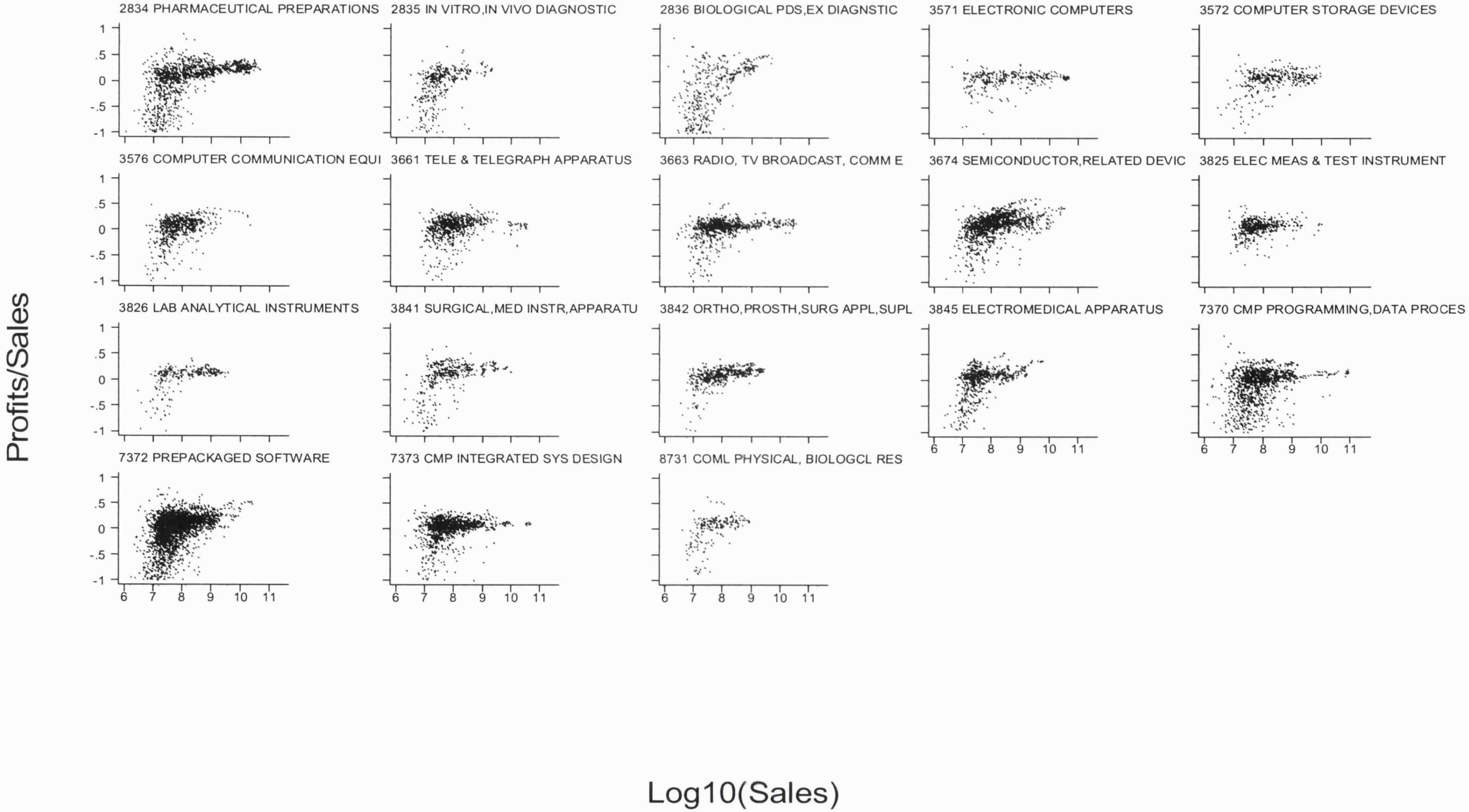
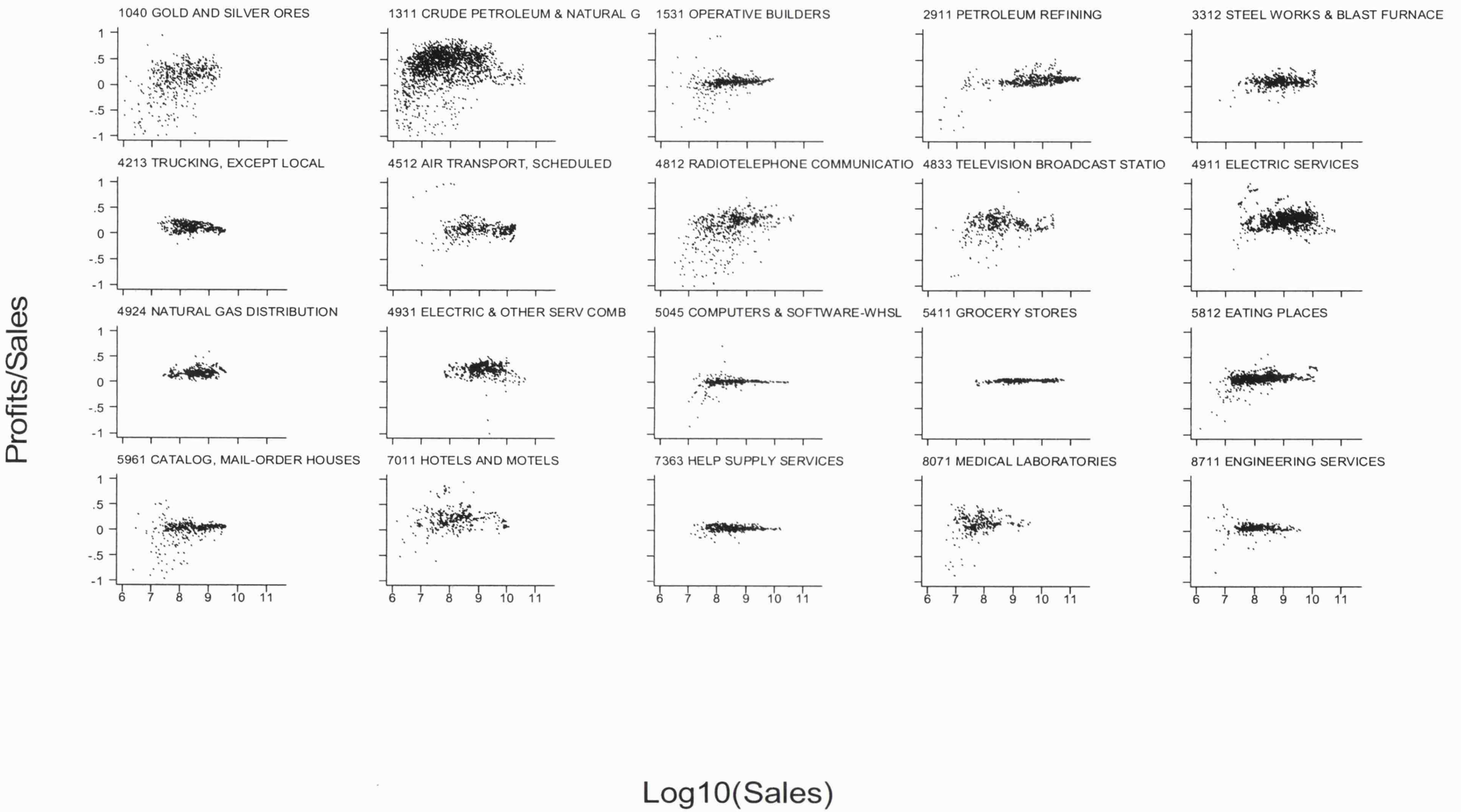


Figure 1-6: Profits/sales against log sales by industry - Low R&D Industries



industries than on low R&D industries.[7] However, a comparison of Figure 1-7 and 1-8 suggests that the general phenomenon of a large number of small loss-making firms in high R&D industries compared to low R&D industries persists.[8]

[7]Missing R&D expenditure data are treated as zeros and the percentage of missing R&D observations is much higher for low R&D industries than for high R&D industries (Table 1.1).

[8]For completeness, Figure 1-9 shows the scatter plots for the group of excluded industries with R&D intensity between 1% and 4%.

Figure 1-7: Gross profits/assets against log assets by industry - High R&D Industries

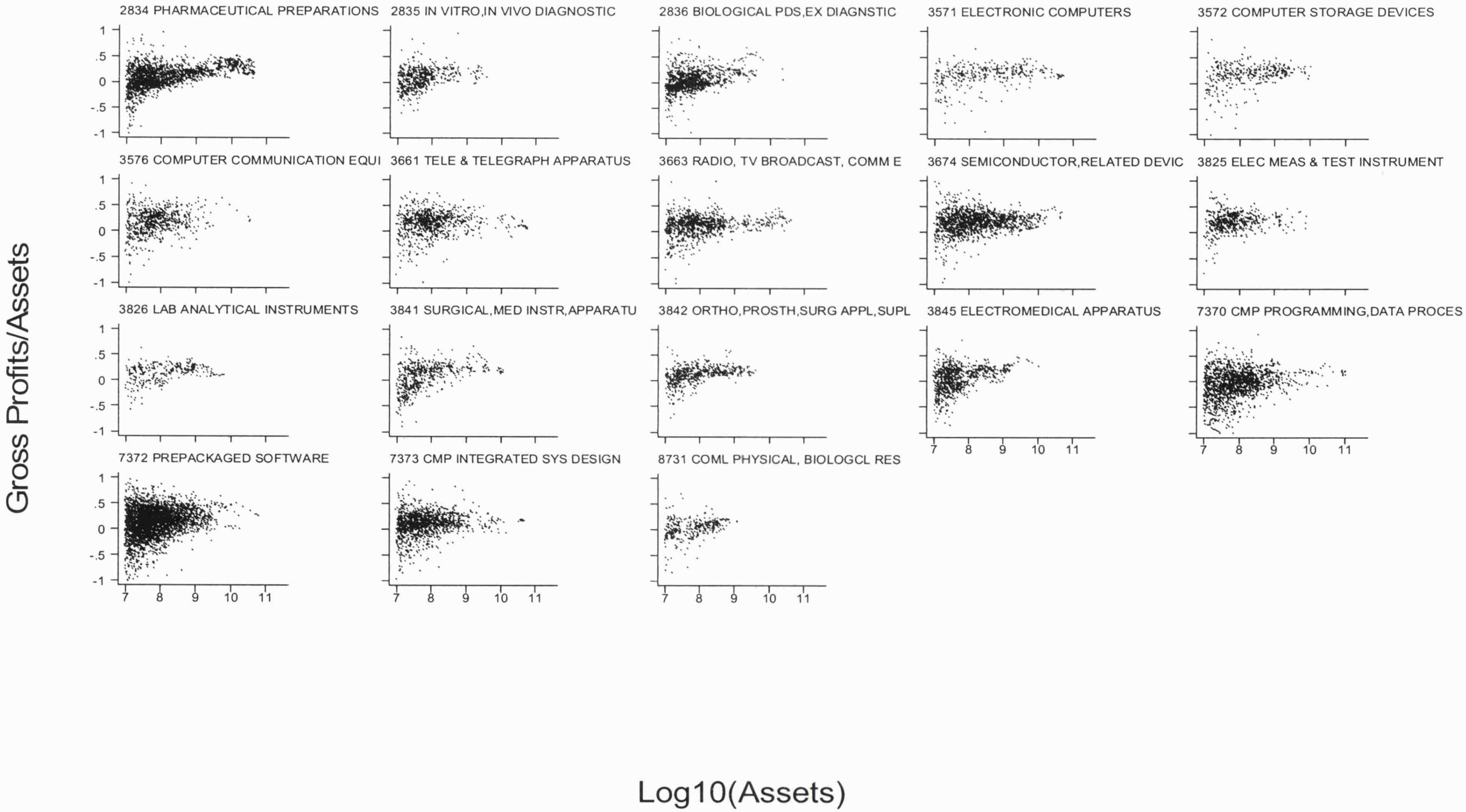


Figure 1-8: Gross profits/sales against log assets by industry - Low R&D Industries

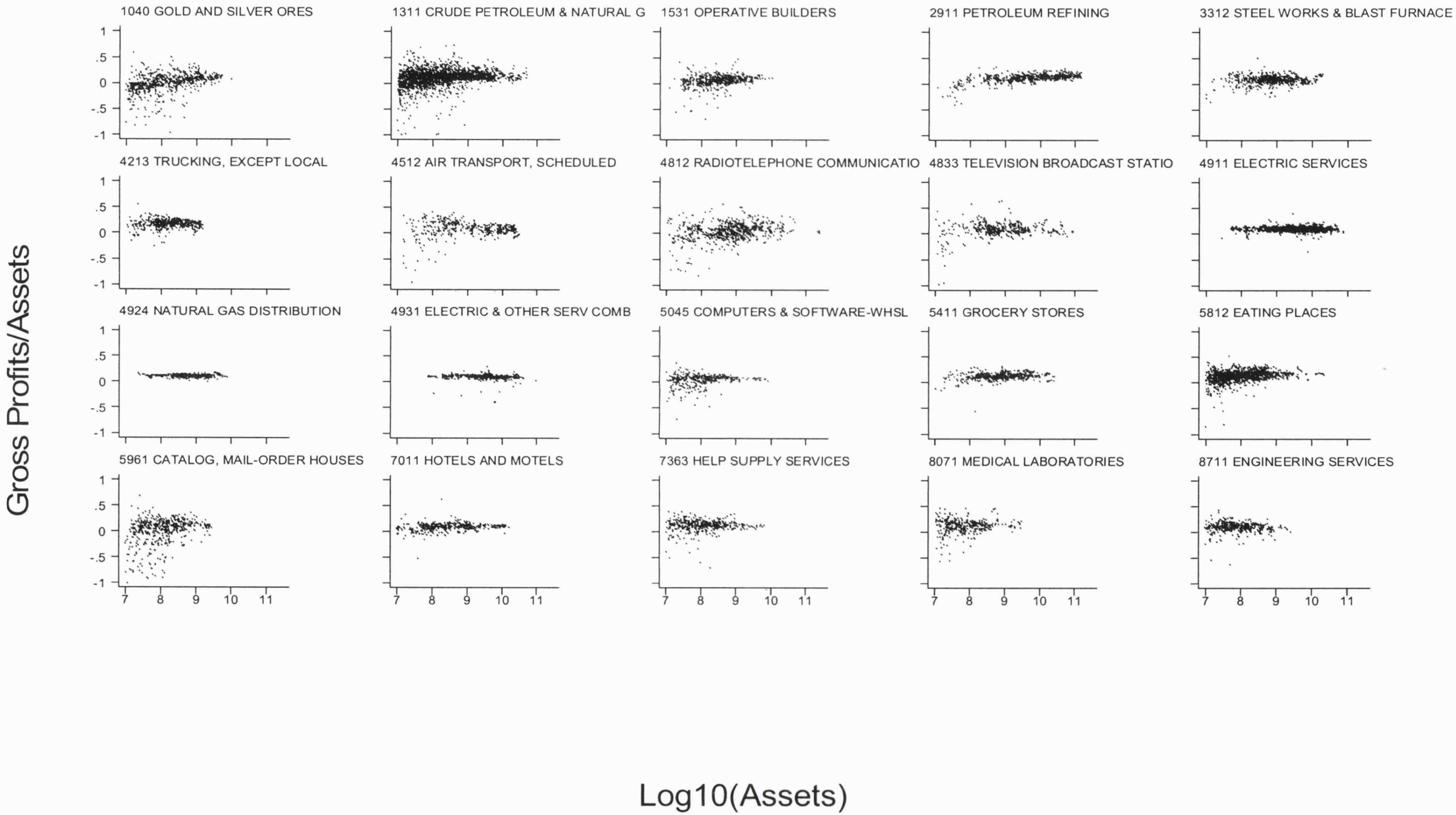
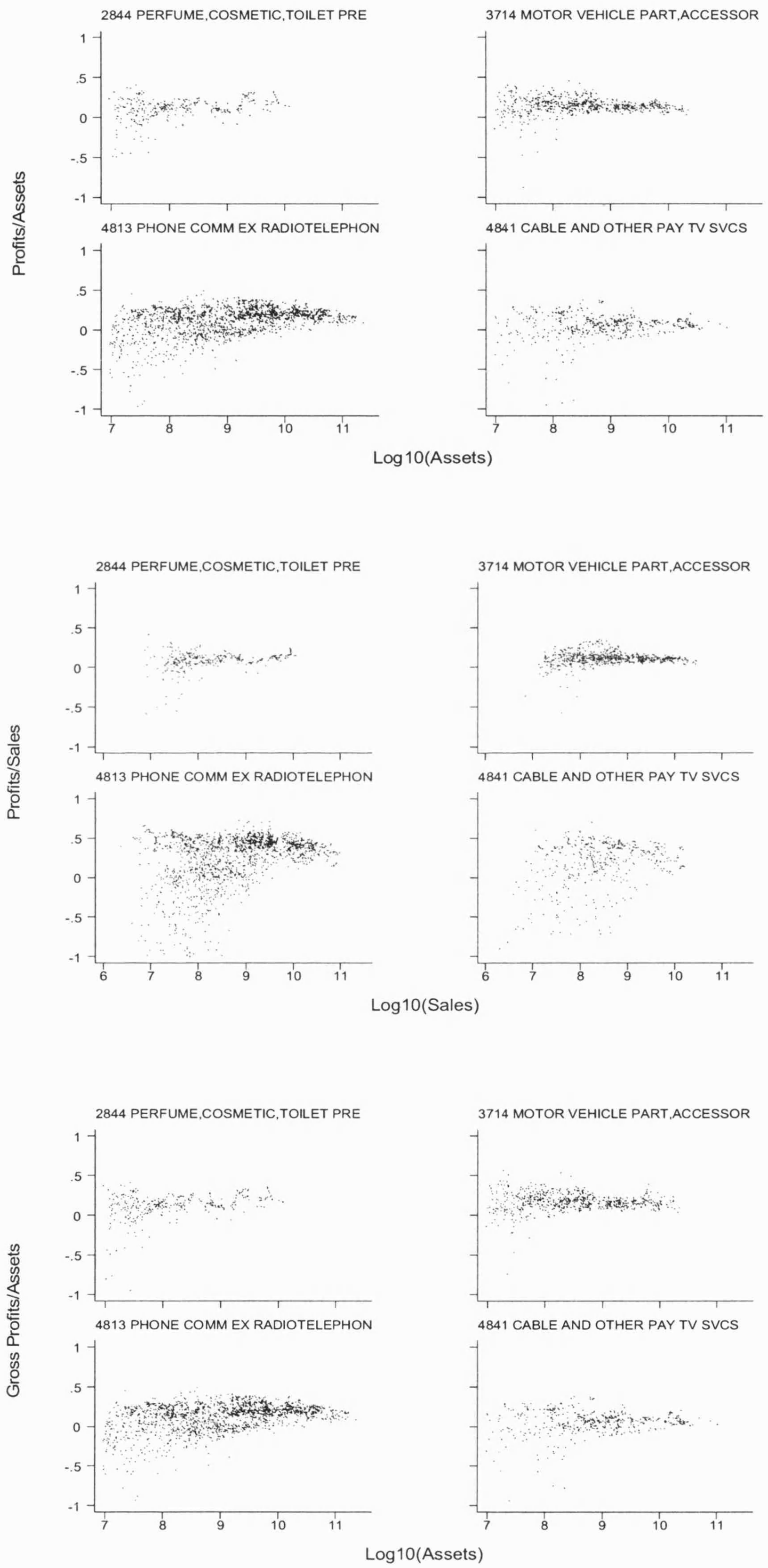


Figure 1-9: Profits-size distribution for excluded medium R&D industries



Estimation of cross industry differences in the distribution

This section characterises the cross industry differences in the profits size distribution more formally using the semiparametric technique of quantile regressions. Quantile regression is a generalisation of median regression (Koenker & Basset(Jr) (1978), Buchinsky (1994)) and has several advantages over OLS: It is robust to outliers in the dependent variable; it is more efficient than OLS for many non-normal error distributions; and it allows the researcher to estimate any conditional quantile of the distribution of the dependent variable, not just the mean.

The last point is of particular interest here, since it implies that quantile regressions can be used to characterise the entire distribution of profitability conditional on size and industry characteristics. That is, rather than making assumptions about the distribution of the dependent variable conditional on observables, quantile regressions provide a location model that allows the researcher to trace out the entire conditional distribution of the dependent variable. It can therefore be used to formally analyse whether the "negative tail" of small loss-making firms in the profits size distribution is systematically stronger in high R&D industries than in low R&D industries.

Assume that the qth quantile of the distribution of the dependent variable y conditional on the independent variables x is given by the linear relationship $x'\beta_q$. The estimated coefficient vector b_q for the conditional quantile q is then the solution to the following programme:

$$b_q = \arg\min_{\beta_q} q \sum_{r_i>0} r_i - (1-q) \sum_{r_i<0} r_i \text{ where } r_i = y_i - x_i'\beta_q$$

Note that with $q = 0.5$ the problem reduces to minimising the sum of the absolute value of the residuals r over the parameter vector β_q, i.e. to a median regression. When q is set to a different value e.g. $q = 0.25$, the objective function is a weighted sum of the absolute deviations giving negative residuals (corresponding to points below the regression line) a weight of $1 - q = 75\%$, and positive residuals (corresponding to points above the regression line) a weight of $q = 25\%$. Giving a higher weight to negative residuals will "force the regression line down" and with

$q = 0.25$, the solution provides an estimate of the 25% quantile of the distribution of the dependent variable conditional on the independent variables.

In the analysis that follows, we present results of quantile regressions of profitability conditional on size and a dummy for industry R&D intensity. However, R&D intensity of the industry is clearly an endogenous outcome of underlying technological primitives that determine the effectiveness of R&D and hence the incentives to invest in R&D. We use R&D intensity at the industry level as a proxy for the differences in the stochastic environment and technology across industries.[9] However, the use of an endogenous variable to proxy for technology calls for a careful interpretation of regression results. Significant coefficients on the industry R&D intensity in the regressions should not be interpreted as a causal relationship between R&D intensity and the conditional distribution as both the profits-size distribution and the R&D intensity are endogenous outcomes of underlying technological primitives. The aim here is simply to document a significant correlation between the conditional distribution of profitability and size and the R&D intensity which proxies differences in technology.

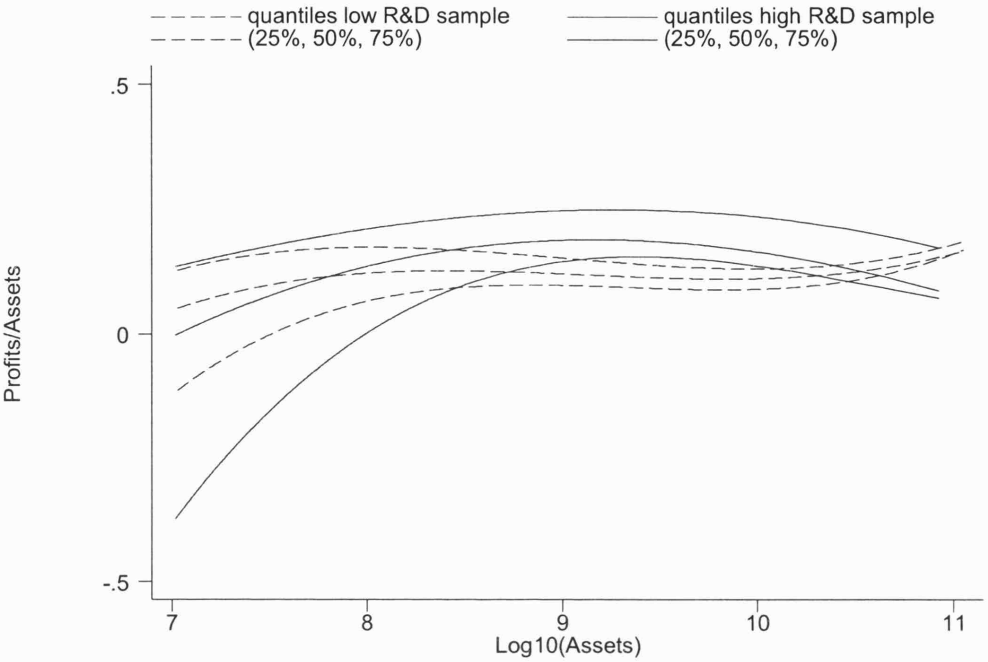
industry effects?

Differences in the conditional regression quantiles across the two subsamples are most easily analysed graphically. Figure 1-10 plots the predicted quantiles for a single year, 1994.[10] The lines shown in the figure are predicted values from quantile regressions of profits/assets on a third order polynomial in log assets. Regressions were performed separately for the high and low R&D subsample on the 25%, 50%, and 75% quantile. The solid lines correspond to the predicted quantiles for the high R&D subsample, while the dashed lines trace the predictions for the low R&D subsample. The figure shows that the predicted 25% quantile and the median are lower for high R&D industries than in low R&D industries for small firms with asset values below US$ 100 million (10^8). The dispersion is also higher for high R&D firms, as the 75% quantile for high R&D industries lies close to, but above

[9] Sutton (1998, Theorem 3.3) proves in the context of a two stage model that in any industry equilibrium configuration a high observed industry R&D intensity implies that the technology parameter has to be such that R&D is effective in influencing future profits. In that sense, differences in the observed industry R&D intensity can proxy for underlying technological differences.

[10] To avoid time effects, the graphs in Figures 1-10 to 1-12 are based on a single year, 1994. The conclusions are insensitive to the choice of year.

Figure 1-10: Quantiles of the profits/assets distribution by R&D intensity, 1994



that for low R&D industries. However, for bigger firms in the range of US$ 1 billion to 10 billion (10^9 to 10^{10}) in assets the effect seems to be reversed as the quantiles for the high R&D subsample lie above those for low R&D industries. This confirms the visual impressions from Figures 1-3 and 1-4: The negative tail of small loss-making firms in the profits size distribution is systematically stronger for high R&D industries than for low R&D industries.

The same conclusion can be drawn from Figure 1-11, which shows the corresponding predicted quantiles for regressions of profits/sales on a polynomial of log sales: The 25% quantile and the median profit rates for small firms are much lower in the high R&D subsample than in the low R&D subsample.

However, a different picture emerges from Figure 1-12, which uses gross profits/assets as measure for profitability. Here, the predicted quantiles for the high R&D subsample are higher than those for the low R&D sample which contrasts the impression from Figures 1-7 and 1-8 that the negative tail is also robust to the choice of this measure.

Table 1.5 reports regression results for another set of quantile regressions. For each regression, there are three columns corresponding to the coefficient vector for

Figure 1-11: Quantiles of the profits/sales distribution by R&D intensity, 1994

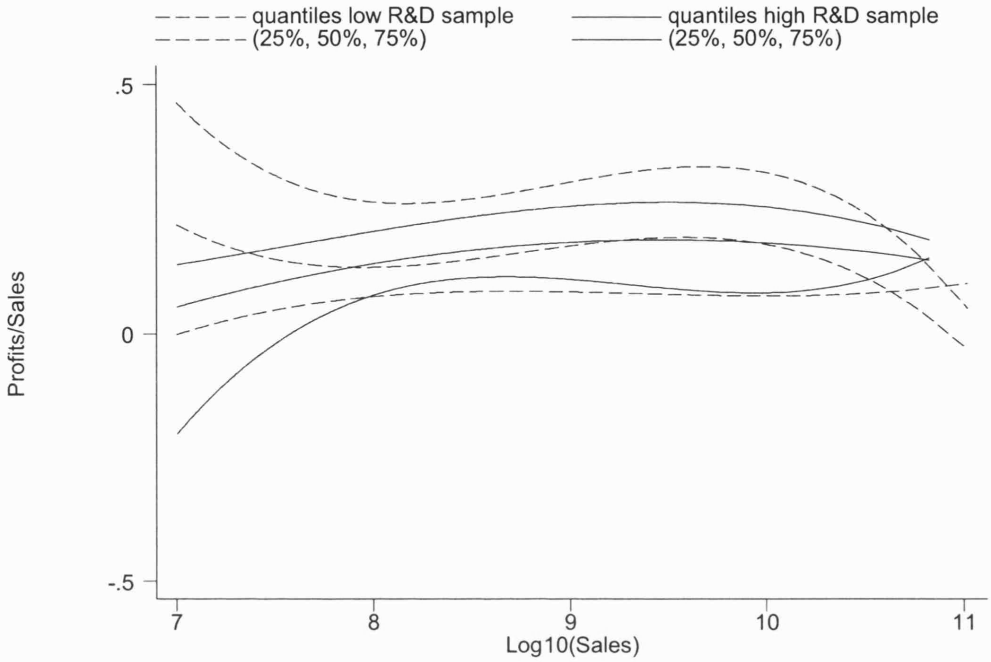
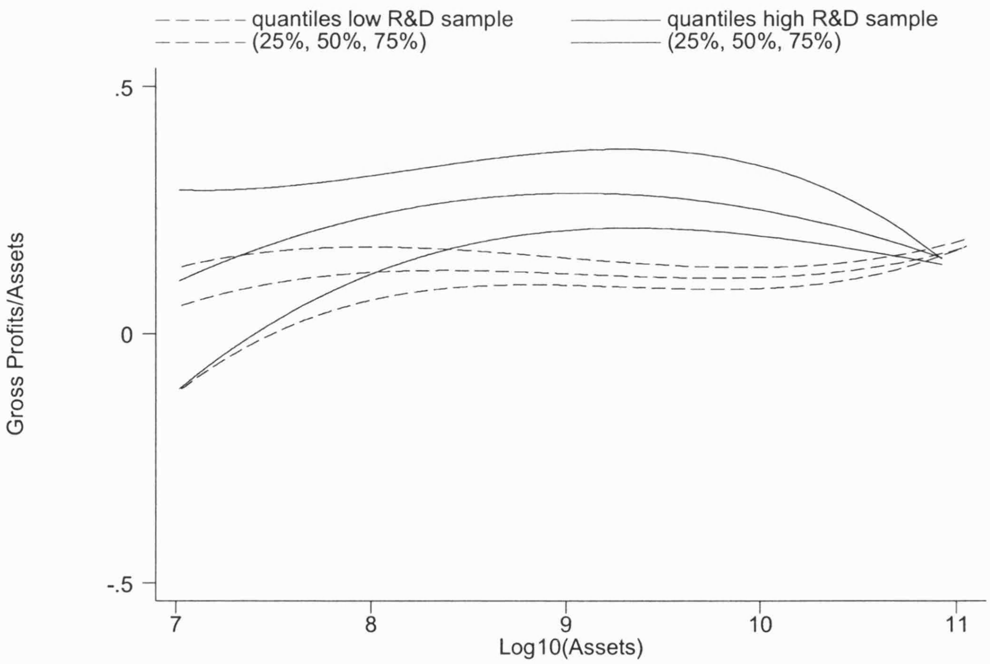


Figure 1-12: Quantiles of the gross profits/assets distribution by R&D intensity, 1994



the 25%, 50%, and 75% quantile. Below the coefficients, bootstrapped standard errors are reported.[11] The regressions were run on the pooled sample from 1990 to 2002 with time dummies. The dependent variable is profits/assets which is regressed on a third order polynomial of log assets and a dummy taking the value one for firms in high R&D industries (Regression 1). The coefficient on R&D for the 25% and 50% quantile is negative, as expected. As the standard errors suggest that these coefficients may not be significant, the p-values from the 50 bootstrap repetitions are also reported. These show, that the hypothesis of a non-negative coefficient for the R&D intensity can be rejected at a confidence level of 2% for the 25% quantile and 8% for the median. For the 75% quantile, the coefficient on R&D is insignificant.

The second regression in the table also includes an interaction term between the R&D dummy and log assets. While the dummy retains a significantly negative coefficient for all quantiles, the interaction term is significantly positive. This confirms, that in high R&D industries, the conditional profitability distribution is shifted down for small firms, but for firms with assets values above 1 billion US$, this effect is reversed.

Finally, Regression 3 also includes a (continuous) measure for the industry advertising intensity to control for this additional industry characteristic. The results with respect to R&D are robust to the inclusion of this measures.[12]

Table 1.6 reports a quantile regression with profits/sales as the dependent variable. The R&D dummy still has the expected negative sign, although the significance is somewhat lower with a p-value of 8% for the 25% quantile and 6% for the median and the 75% quantile.

Table 1.7 reports the same regression with gross profits/assets as the dependent variable. As suggested by Figure 1-12, the negative tail in high R&D industries fails to emerge with this profits measure and the median and 75% quantile of profits

[11] These are based on 50 repetitions. The bootstrapped samples are drawn with replacement from the set of 38 4-digit SIC industries. The unit of observation is the industry and each industry is assigned equal probability to be selected. For each sample, the random draws were stopped when the number of firm-year observations in the sample equalled or just exceeded the number of observations in the original sample.

[12] The regressions also proved robust to the use of the industry R&D intensity as a continuous measure and the inclusion of the a measure of capital intensity (not reported).

Table 1.5: Quantile regressions for profits/assets

Dep. Var: Profits/assets	Regression 1			Regression 2			Regression 3		
Quantile	25	50	75	25	50	75	25	50	75
log(assets)	**8.926**	**4.993**	**2.792**	**6.209**	**4.075**	**2.235**	**6.463**	**4.055**	**2.155**
SE	1.183	0.950	0.679	1.177	0.929	0.625	1.146	0.953	0.609
log^2(assets)	**-0.942**	**-0.534**	**-0.302**	**-0.657**	**-0.442**	**-0.249**	**-0.687**	**-0.440**	**-0.239**
SE	0.124	0.103	0.078	0.128	0.103	0.072	0.124	0.106	0.070
log^3(assets)	**0.033**	**0.019**	**0.011**	**0.023**	**0.016**	**0.009**	**0.024**	**0.016**	**0.009**
SE	0.004	0.004	0.003	0.005	0.004	0.003	0.004	0.004	0.003
R&D-dummy	**-0.080**	**-0.022**	**0.009**	**-1.056**	**-0.604**	**-0.436**	**-1.073**	**-0.611**	**-0.408**
SE	0.050	0.022	0.013	0.272	0.209	0.099	0.269	0.204	0.100
P-value	0.020	0.080	0.360						
R&D-dummy				**0.114**	**0.069**	**0.053**	**0.117**	**0.070**	**0.050**
* log(assets)				0.029	0.023	0.011	0.029	0.023	0.012
Adv/Sales							**-1.112**	**-0.101**	**0.428**
							1.292	0.941	0.697
pseudo R-squared	0.208	0.075	0.021	0.227	0.088	0.033	0.229	0.088	0.034
# obs	32247			32247			32247		
# industries	38			38			38		

All regressions include time dummies and an intercept and cover the years 1990 to 2002.
Bootstrapped standard errors are reported below the coefficients and are obtained using 50 repetitions treating the industry as a unit of observation and drawing industries with replacement and with equal probability to be selected until the bootstrap samples contained as many or just over the number of firm-year observations as the original sample.

of high R&D industries in this specification lie higher than in low R&D industries.

Overall, the graphs on the industry distributions and the quantile regressions give strong evidence that the negative tail in the joint distribution for profits and size is longer in high R&D industries than in low R&D industries for the profitability measures profits/assets and profits/sales, but not for gross profits/assets.

why?

Table 1.6: Quantile regressions for profits/sales

Dep. var: Profits/sales

Quantile		25	50	75
log(sales)		**223.184**	**94.448**	**42.486**
	SE	57.007	26.247	14.508
log^2(sales)		**-25.687**	**-10.949**	**-4.950**
	SE	6.836	3.152	1.733
log^3(sales)		**0.979**	**0.421**	**0.191**
	SE	0.271	0.126	0.069
RD-dummy		**-0.113**	**-0.124**	**-0.157**
	SE	0.078	0.089	0.101
	Pval	0.082	0.060	0.060
pseudo R-squared		0.155	0.072	0.033
# obs		31659		
# industries		38		

The regressions include time dummies and cover the years 1990 to 2002. Bootstrapped standard errors (based on 50 repetitions treating the industry as the unit of observation) are reported below the coefficients.

Table 1.7: Quantile regressions for gross profits/assets

Dep. var: Gross profits/assets

Quantile		25	50	75
log(assets)		**5.200**	**4.307**	**2.922**
	SE	0.760	0.639	0.617
log^2(assets)		**-0.546**	**-0.464**	**-0.321**
	SE	0.083	0.070	0.070
log^3(assets)		**0.019**	**0.017**	**0.012**
	SE	0.003	0.003	0.003
RD-dummy		**0.025**	**0.067**	**0.115**
	SE	0.031	0.024	0.020
	Pval	0.240	0.000	0.000
pseudo R-squared		0.121	0.047	0.078
# obs		32247		
# industries		38		

The regressions include time dummies and cover the years 1990 to 2002. Bootstrapped standard errors (based on 50 repetitions treating the industry as the unit of observation) are reported below the coefficients.

1.2.3 Intra distribution dynamics

Differences in the profits-size distribution will, of course, be endogenously determined by differences in technological and stochastic primitives driving the underlying firm dynamics with respect to profitability and size. The cross sectional distributions are only a snapshot of the current profitability and size of firms, whereas firms will base their decisions, including entry and exit decisions, on their expectation of the future evolution of their profit flow. Sustaining losses without exiting will only be justified when the expected net present value of future profit flows is positive. The dynamics of profits and size over time are, therefore, crucial in determining exit decisions and hence the shape of the cross sectional distribution.

In this section, we will develop some stylised facts on differences in the intra distribution dynamics across high and low R&D industries. We first examine the variation of profit rates and size from one year to the next. Then stylised facts of the evolution of firms over a given time horizon are derived by means of transition matrices. To do so, the support of the profits-size distribution is divided in three discrete states and we examine transition probabilities between these states and exit states. The stylised facts on intra distribution dynamics derived in this way will motivate modelling choices in the second part of the chapter and will provide a qualitative benchmark for the model.

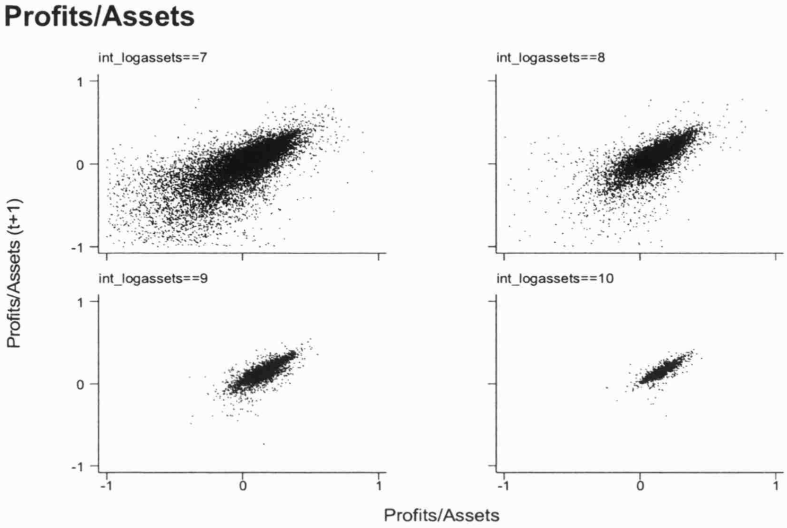
π_{t+1} π_t why not?

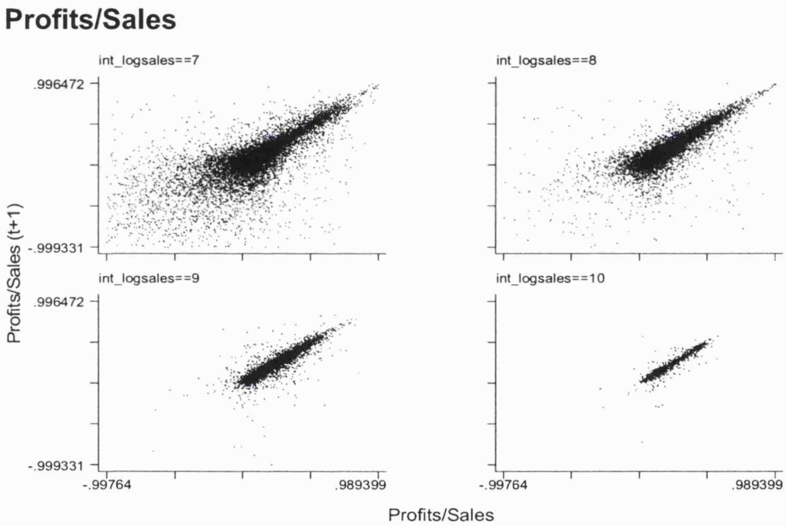
Year on year changes in profit rates

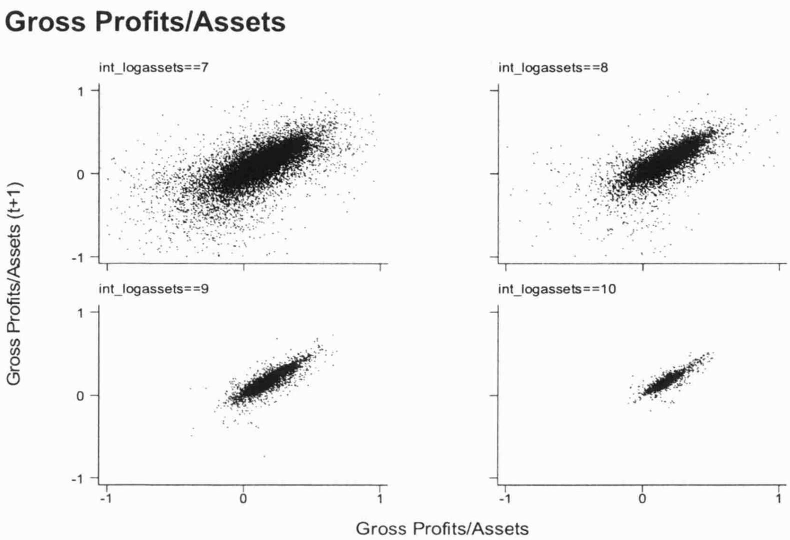
As a first cut on the intra distribution dynamics, we examine changes in profit rates. Figure 1-13 panel a (respectively panels b and c) shows scatter plots of profits/assets (respectively profits/sales and gross profits/assets) of firms in year t against profits over assets (sales) in year $t+1$ for 4 size classes corresponding to the integer of their base 10 logarithm of assets (sales) (i.e. up to US$ 100 million; 100 million -1 billion; 1-10 billion and above 10 billion). The fact that the points are clustered along the diagonal implies that there profit rates are persistent from year to year. The pictures also suggests that the dispersion of the change in profit rates, conditional on firm survival, decreases with the current profit rate and with firm size. The latter point is consistent with a diversification argument.

To confirm this impression and to investigate differences across high and low

Figure 1-13: Year on year change in profitability by size class







R&D industries, we run another set of quantile regressions. We estimate the interquartile range of the change in profitability (i.e. the first difference in profit rates) in for each of the profitability measure (Table 1.8). The interquartile range is the difference between the 75% quantile and the 25% quantile and measures the dispersion of the change in profits rates. For the profits/assets regression the variance in the change of profit rates is significantly higher in high R&D industries. This is also the case for the gross profits/assets regression. However, the coefficient of R&D is insignificant and of the wrong sign for the profits/sales regression. This may be due to the very high dispersion in profit/sales measures for the "Crude petroleum and natural gas (SIC 1311)" industry which, with over 2000 observations is the biggest low R&D industry in our sample.

Apart from the profits/sales specification, it therefore seems that the (conditional) dispersion of the change in profit rates is higher for high R&D firms, supporting the view that the variance of future profits is higher in high R&D industries than in low R&D industries.

Transitions between discrete states in the distribution

A second approach to assess differences in the intra distribution dynamics is to discretise the support of the joint profits-size distribution into a number of discrete states and to examine differences across industries in transition probabilities between these states and exit states over a given time horizon. We use two different discretisations. Each divides the support of the profits-size distribution in three states. In both discretisations, one state contains big firms which are defined as firms with asset values exceeding US$ 1 billion. The two approaches differ with respect to the definition of states for small firms. The first distinguishes between small profitable and small unprofitable firms, i.e. is based on an absolute threshold of zero profits. The second approach divides the set of small firms in firms with profits/assets above the median small firm in the industry and firms below the median firm, i.e. uses the position relative to the median small firm in the industry. We present transition probability matrices for the group of high and low R&D firms for each approach below for transitions over the years 1994 to 1998. The rows of each matrix correspond to the state in 1994 and the columns to the

Table 1.8: Quantile regressions on the change in profit rates

Profits/Assets(t+1) - Profits/assets(t)		Profits/sales(t+1) - Profits/sales(t)		Gross P/A(t+1) - Gross P/A(t)	
Profits/Assets	**-0.313**	**Profits/sales**	**-0.317**	**Gross P/A**	**-0.087**
	0.027		0.091		0.029
log(assets)	**-0.021**	**log(sales)**	**-0.035**	**log(assets)**	**-0.024**
	0.003		0.043		0.005
R&D-dummy	**0.060**	**R&D-dummy**	**-0.019**	**R&D-dummy**	**0.067**
	0.008		0.065		0.009
Const.	**0.556**	**Const.**	**0.556**	**Const.**	**0.270**
	0.032		0.401		0.046
# obs	26915	**# obs**	26377	**# obs**	26915
# industries	38	**# industries**	38	**# industries**	38
pR2 .75	0.166	**pR2 .75**	0.622	**pR2 .75**	0.120
pR2 .25	0.041	**pR2 .25**	0.071	**pR2 .25**	0.073

All regressions include an intercept and time dummies and cover the years 1990 to 2001.
Bootstrapped standard errors are reported below the coefficients and are obtained using 50 repetitions treating the industry as a unit of observation and drawing industries with replacement and with equal probability to be selected until the bootstrap samples contained as many or just over the number of observations as the original sample.

state in 1998:

"Absolute" states: Table 1.9 (panels a and b) report transition probability matrices for high and low R&D industries on the basis of an absolute threshold of zero profits between the two states for small firms. The following stylised facts emerge:

- **Small profitable firms in high R&D industries** have
  - a higher probability of becoming unprofitable (15.5%) than low R&D firms (8.5%) and
  - a lower probability of going bankrupt or exit through "other reasons" (2.2%) over the 4 year horizon than low R&D firms (5.5%).
- **Small unprofitable firms in high R&D industries** have
  - a much higher probability of remaining in this state (50.9%) than firms in low R&D industries (30.7%) and
  - a lower probability of going bankrupt or to exit through "other reasons" (8.7%) than low R&D firms (23.8%).

An analysis of the transition matrices industry by industry (not reported here) shows, however, that within the group of high R&D industries, "Pharmaceutical" and "Biotech" industries (SIC's 2384-2386) are outliers in the sense that they have huge numbers of firm in the small unprofitable state and that these firms have particularly high probabilities of staying in this state. Therefore, we present in panel c the transition matrices for high R&D industries excluding "Pharmaceuticals" and "Biotech":

- **small unprofitable firms in high R&D industries excluding "Pharmaceuticals" and "Biotech"** have
  - a higher (but much closer) probability of remaining unprofitable,
  - a higher probability of becoming profitable, and
  - a lower probability of exiting through bankruptcy/liquidation or "other reasons" than low R&D firms in the same state.

Table 1.9: Transition probability matrices for ABSOLUTE states

(Transition probabilities with SE's below - both in %)

Panel a: LOW R&D industries

ABSOLUTE STATE in 1994		ABSOLUTE STATE IN 1998: large	small prof.	small unprof.	M&A	Bkr/Lqn	Other exit	Total exit	#firms
large	TP	83.9	2.8	0.0	12.4	0.0	0.8	13.3	354
	SE	2.0	0.9	0.0	1.8	0.0	0.5	1.8	
small prof.	TP	7.1	57.0	8.5	21.9	1.9	3.6	27.4	686
	SE	1.0	1.9	1.1	1.6	0.5	0.7	1.7	
small unprof.	TP	2.0	19.8	30.7	23.8	10.9	12.9	47.5	101
	SE	1.4	4.0	4.6	4.2	3.1	3.3	5.0	
Total		30.5	36.9	7.8	19.1	2.1	3.6	24.8	1141

Panel b: HIGH R&D industries

ABSOLUTE STATE in 1994		ABSOLUTE STATE IN 1998: large	small prof.	small unprof.	M&A	Bkr/Lqn	Other exit	Total exit	#firms
large	TP	84.1	1.2	0.0	14.6	0.0	0.0	14.6	82
	SE	4.0	1.2	0.0	3.9	0.0	0.0	3.9	
small prof.	TP	6.5	51.0	15.5	24.8	1.4	0.8	27.0	718
	SE	0.9	1.9	1.3	1.6	0.4	0.3	1.7	
small unprof.	TP	1.0	20.8	50.9	18.7	2.1	6.6	27.3	289
	SE	0.6	2.4	2.9	2.3	0.8	1.5	2.6	
Total		10.9	39.2	23.7	22.4	1.5	2.3	26.2	1089

Panel c: HIGH R&D industries excluding "Pharmaceuticals" and "Biotech" industries - SIC 2834-2836

ABSOLUTE STATE in 1994		ABSOLUTE STATE in 1998: large	small prof.	small unprof.	M&A	Bkr/Lqn	Other exit	Total exit	#firms
large	TP	85.2	0.0	0.0	14.8	0.0	0.0	14.8	54
	SE	4.8	0.0	0.0	4.8	0.0	0.0	4.8	
small prof.	TP	5.9	49.8	16.2	25.7	1.5	0.9	28.1	647
	SE	0.9	2.0	1.4	1.7	0.5	0.4	1.8	
small unprof.	TP	1.3	28.8	37.8	19.9	3.2	9.0	32.1	156
	SE	0.9	3.6	3.9	3.2	1.4	2.3	3.7	
Total		10.0	42.8	19.1	23.9	1.8	2.3	28.0	857

large: assets>US$ 1 billion
small, profitable: profits>0 and assets<=US$ 1 billion
small, unprofitable: profits<=0 and assets<=US$ 1 billion

"Relative" states: Defining the states for small firms on the basis on the profitability of the median small firm in the industry is a way of controlling for industry effects. (When using this relative state definition, excluding "Pharmaceuticals and Biotech" does not change the transition matrix for high R&D firms significantly). The two states "small, high (relative) profits" and "small, low (relative) profits" allow us to investigate how firms evolve within their relative distribution, i.e. relative to their peers. With this relative state definition, it seems that (Table 1.10):

- Transition probabilities for firms in the "**small, high relative profits**" state and for "**big**" firms are fairly similar for high and low R&D industries.
- Firms in the "**small, low relative profits**" state in high R&D industries are still
  - more likely to survive in this state and are
  - less likely to go bankrupt.

  However, the differences across the two subsamples are much less pronounced than with "absolute" states.

Table 1.10: Transition probability matrices for RELATIVE states

(Transition probabilities with SE's below - both in %)

Panel a: LOW R&D industries

RELATIVE STATE in 1994		RELATIVE STATE IN 1998: large	small hi prof.	small lo prof.	M&A	Bkr/Lqn	Other exit	Total exit	#firms
large	TP	83.9	2.0	0.8	12.4	0.0	0.8	13.3	354
	SE	2.0	0.7	0.5	1.8	0.0	0.5	1.8	
small hi prof.	TP	6.9	40.8	23.0	24.8	1.5	3.0	29.2	404
	SE	1.3	2.4	2.1	2.1	0.6	0.8	2.3	
small lo prof.	TP	6.0	19.3	43.9	19.3	4.7	6.8	30.8	383
	SE	1.2	2.0	2.5	2.0	1.1	1.3	2.4	
Total		30.5	21.6	23.1	19.1	2.1	3.6	24.8	1141

Panel b: HIGH R&D industries

RELATIVE STATE in 1994		RELATIVE STATE IN 1998: large	small hi prof.	small lo prof.	M&A	Bkr/Lqn	Other exit	Total exit	#firms
large	TP	84.1	1.2	0.0	14.6	0.0	0.0	14.6	82
	SE	4.0	1.2	0.0	3.9	0.0	0.0	3.9	
small hi prof.	TP	7.4	42.8	23.5	24.3	1.3	0.8	26.3	638
	SE	1.0	2.0	1.7	1.7	0.4	0.3	1.7	
small lo prof.	TP	0.8	20.3	50.4	20.9	2.2	5.4	28.5	369
	SE	0.5	2.1	2.6	2.1	0.8	1.2	2.3	
Total		10.9	32.0	30.9	22.4	1.5	2.3	26.2	1089

large: assets>US$ 1 billion
small, high profit: profits/assets>median profits/assets and assets<=US$ 1 billion
small, low profit.: profits/assets<=median profits/assets and assets<=US$ 1 billion

1.3 A simple stochastic model for firm dynamics

The systematic variation of the profits-size distribution and the firm dynamics with industry R&D intensity indicates that there are underlying driving mechanisms that are common across industries. The key theoretical idea of the chapter is that the observed empirical regularities are due to a combination of two effects: a real option effect explaining the existence and survival of loss-making firms; and a diversification effect leading to the decrease in the cross sectional variance of profit rates with firm size. In this section, we propose a theoretical reduced form model for firm dynamics that combines these two effects by modelling firms as consisting of (approximately) independent businesses. We then investigate whether this model can generate the empirical regularities in the first part of the chapter by means of simulations.

The essence of the real option effect is as follows: When exit decisions are definite (i.e. a firm cannot temporarily suspend its operations), loss-making firms will weigh current losses against the expected net present value of future profits (net of opportunity costs). With stochastic future profit flows, there is an option value of not exiting: Rather than exiting today and thereby limiting losses, the firm can always wait for the next profit realisation and then make its decision on the basis of this updated information. Waiting provides the firm with the option to enjoy potentially high future profits should the profit flow develop favourably, while still leaving the option of exit (the ability to cut losses) if things get worse. The firm will weigh this option value of not exiting against the cost of waiting, i.e. the current losses. Only when the (flow of) current losses exceeds the option value of waiting, will the firm optimally decide to exit. This effect can explain why many loss-making firms do not exit immediately. It can hence generate a tail of loss-making firms in the profits-size distribution.

The crucial insight for generating the systematic differences in the length of the negative tail across high and low R&D industries is that, conditional on the current profit flow, the option value of waiting increases with the variance of the future profit flow. This is because the upside risk of high future profits increases with the variance of future profits, while the firm can always limit its downside risk of low future profits by exiting. As a result, firms with a higher variance in future profits are willing to suffer higher current losses before they optimally decide to exit. This provides a rationale for why the negative tail in the profits size distribution is stronger in high R&D than in low R&D industries: If R&D intensity is a proxy for a higher variance or higher uncertainty about future profits, firms in high R&D industries will find it optimal to stay in the market at negative profit levels at which their counterparts in low R&D industries will long have decided to exit.

To generate a rationale for a decline in the cross sectional variance in firm profitability with firm size, the diversification effect will come into play. For firms consisting of a number of businesses, aggregate profitability will be a weighted average across the profitability of the constituent businesses. If bigger firms tend to consist of a larger number of businesses, and if these businesses evolve independently from each other, a law of large numbers effect will drive the decline in the cross sectional variance of aggregate firm profitability with firm size.

While independence across the firm's businesses is a strong assumption, it is a useful limiting case that provides a benchmark for evaluating the performance of any model where firms consist of business units that has proved to be very successful in recent work (Sutton (1998, chapter11), Sutton (2002)).

1.3.1 Description of the model

Consider a single agent model for a risk neutral firm in continuous time. The firm consists of one or more businesses. New business opportunities arrive randomly over time. At any point in time, the firm can decide to shut down each of its businesses. Once shut down, a business opportunity cannot be reopened, so that temporary suspension of the business is ruled out. The firm exits, when it abandons the last of its businesses.

The profit flow from business i is a function of this business' state ω_i, which evolves stochastically over time. The only decision available to the firm with respect to each business is whether to continue the business or to abandon it.[13] The abandonment decision is costless and final and results in a zero payoff.

[13]One interpretation is that this is a reduced form of a model where all other decision variables have already been optimised out.

Φ < R and D intens non-tangible
Non R and D can sell asset

To keep the model as simple as possible, two additional assumptions are introduced. These assumptions will imply additive separability of the firm's value function across business units and independence of the shut down decisions of the number and states of the active businesses. The first assumptions is that the state of each business evolves independently from the state of all other businesses. This ensures that the shut down decision for a business is a function of this business' state only and does not depend on the states of the firm's other active businesses.

The aim of the second assumption is to ensure that the exit decision for the *firm* (shutting down the last business) is no different from the decision to shut down any other *business*. To achieve this, assume that there are no fixed costs of operation at the firm level and that the expected value of future businesses Γ is equal to the value of the firm's outside option Φ. One interpretation of this assumption is that the entrepreneur can always costlessly re-enter with a new firm should new business opportunities arise.

With these assumptions, the firm's value function can be written as the sum of the value of its businesses,

not separate, since?

$$V(\omega) = \max\{\Phi, \sum_{i \in I} v(\omega_i) + \Gamma\}, \tag{1.1}$$

where ω denotes the state vector of the firm, i.e. a vector containing the states of all the firm's active businesses, I denotes the set of the firm's active businesses, and $v(\omega_i)$ is the value function for business i. As discussed above, the expected value of the firm's future business opportunities Γ is assumed to equal the value of the firm's outside option Φ for simplicity.[14]

The value function for each of the constituent businesses is given by the Bellman equation

$$v(\omega_i) = \max\{0, \pi(\omega_i)dt + \frac{1}{1+\rho dt} E[v(\omega_i + d\omega_i)|\omega_i]\}. \tag{1.2}$$

The firm can guarantee itself a profit flow of zero by abandoning the business. Continuing the businesses over the short time interval dt results in a profit flow of

[14] The value of future businesses is a constant Γ and is independent of firm's state ω. Below, this will be achieved by assuming that new businesses arrive with a constant arrival rate λ and with initial states drawn from an iid distribution.

$\pi(\omega_i)dt$ and an expected value of the business after the time interval dt of $E[v(\omega_i + d\omega_i)|\omega_i]$, appropriately discounted by $(1+\rho dt)^{-1}$ where ρ denotes the discount rate. The expectation is taken over the stochastic increment of the state variable $d\omega_i$ over the time interval dt conditional on the current state ω_i and conditional on making optimal decisions in the future (the Bellman principle). The firm will continue the process if the second argument of the right hand side of (1.2) exceeds the first.

Under mild regularity restrictions, equation (1.2) satisfies Blackwell's conditions for a contraction mapping implying that there is a unique solution to the value function in (1.2). The solution will be the value function as well as a shut down rule for the business of the form

$$\chi(\omega_i) = \begin{cases} 0 \text{ (shut down)} & \text{if } \omega_i \in \Xi, \\ 1 \text{ (continue)} & \text{otherwise.} \end{cases}$$

Abandonment of the business is optimal if and only if the state of the businesses ω_i falls in the shut down region Ξ which is a subset of the state space Ω. In the following, the boundary between the continuation region $\Omega \backslash \Xi$ and the shut down region Ξ will be denoted ω^*. Once a solution to the value function and the shut down rule $\chi(\omega_i)$ is found, it can be combined with an arrival process for new businesses to generate profits-size distributions and intra-distribution dynamics through simulations.

Analytical solutions for the value function and the shut down rule will only be available for some combinations of the stochastic specification of the evolution of the state variable ω_i and the instantaneous profit function $\pi(\omega_i)$. To find analytical solutions, two additional conditions which hold at the boundary ω^* between the continuation region and the stopping region are useful:

$$v(\omega^*) = 0 \text{ and } \frac{\partial v(\omega^*)}{\partial \omega} = 0 \tag{1.3}$$

The first is the value matching condition: At the optimal stopping point, the value of the business equals the value of stopping (zero). If the value of the business were positive (negative) at the optimal stopping point, it would be optimal to stop the process at a lower (higher) value of the instantaneous profit flow. The second is a higher order "smooth pasting" condition which says that the value function and the value of stopping must meet tangentially at the optimal stopping point. Similarly, it can easily be shown – for the class of stochastic processes we will consider here – that it is either worthwhile waiting a little longer or stopping earlier if the smooth pasting condition is not satisfied (Dixit & Pindyck 1994, p 130). The argument is that if the condition is violated, the expected payoff of waiting an infinitesimal time interval dt exceeds the abandonment payoff so that shut down cannot be optimal.

Equipped with these general considerations, we will now consider two simple specification for the evolution of the state variable and the instantaneous profit function.

1.3.2 The simplest specification: The profit flow follows a Brownian motion

Consider first the simplest specification, where ω_i is one dimensional and $\pi(\omega_i) = \omega_i$. With a one dimensional state, the boundary ω^* between the continuation region and the shut down region becomes simply a threshold value. If the state falls below this threshold value, it is optimal to shut down the business, i.e.

$$\chi(\omega_i) = \begin{cases} 0 \text{ (shut down)} & \text{if } \omega_i \leq \omega^*, \\ 1 \text{ (continue)} & \text{otherwise.} \end{cases}$$

Because of this property, the threshold value ω^* will be called the "optimal stopping point" for the state of the business. When the process for ω_i hits this value, it is stopped.

Assume that ω_i follows a Brownian motion with drift and variance parameter μ and σ:

$$d\omega_i = \mu dt + \sigma dz,$$

where dz is the increment of a standard Wiener process, i.e. $E(dz) = 0$, $E(dz^2) = dt$. In the continuation region, the Bellman equation (1.2) then becomes:

$$v(\omega_i) = \omega_i dt + \frac{1}{1+\rho dt} E[v(\omega_i + d\omega_i)|\omega_i].$$

Expanding terms, using Ito's lemma, rearranging, dividing by dt, and letting $dt \to 0$ yields the second order differential equation (see Appendix A1.2), where primes denote derivatives:

$$-\rho v(\omega_i) + \mu v'(\omega_i) + \frac{1}{2}\sigma^2 v''(\omega_i) = -\omega_i. \tag{1.4}$$

Imposing the value matching and the smooth pasting condition (1.3) and the additional condition $0 < \lim_{\omega_i \to \infty} v'(\omega_i) < \infty$[15], yields the solution for $v(\omega_i)$ and ω^* (see appendix A1.2):

what role does Φ play

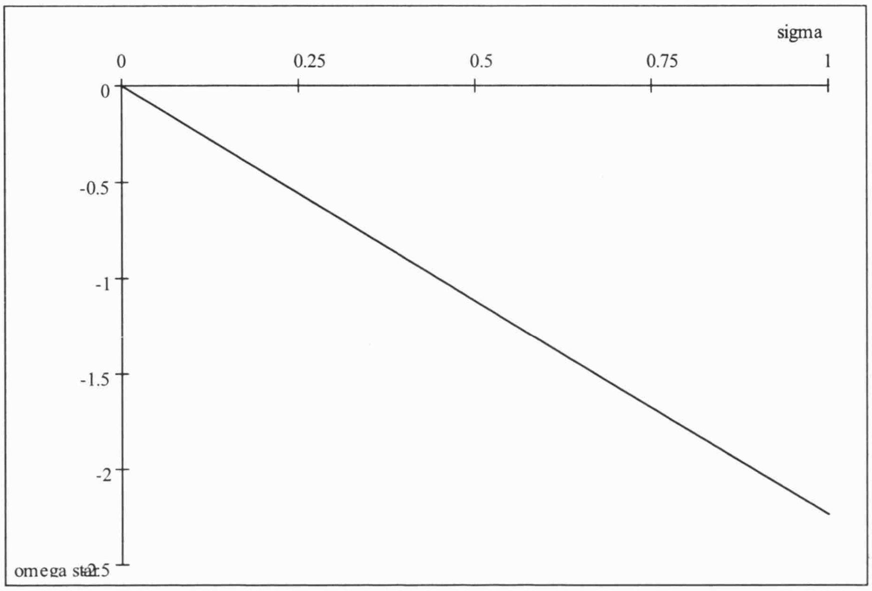
$$\begin{aligned} v(\omega_i) &= -\frac{1}{r\rho}e^{r(\omega_i - \omega^*)} + \frac{1}{\rho}\omega_i + \frac{\mu}{\rho^2} \\ \text{with } r &= -\frac{\mu + \sqrt{\mu^2 + 2\sigma^2\rho}}{\sigma^2} < 0, \text{ and} \\ \omega^* &= -\frac{\sigma^2}{\mu + \sqrt{\mu^2 + 2\sigma^2\rho}} - \frac{\mu}{\rho} < 0. \end{aligned} \tag{1.5}$$

The main interest here lies in the threshold value ω^* for the shut down rule as a function of the variance parameter σ. In this simple case where the state variable is also the profit flow, ω^* is the profit level at which the firm would optimally abandon the business. Hence, ω^* implies "how bad things can get" before the firm optimally decides to abandon the business and gives a lower bound on the support of the profits distribution of a single project. It is easy to show that the exit threshold ω^* is indeed decreasing in the variance parameter σ.[16] Figure 1-14 plots the optimal stopping point ω^* as a function of the variance parameter σ of the underlying process for $\mu = 0$ and $\rho = .1$. As expected, the option value effect embedded in the dynamic programme for the business can therefore explain, why the negative tail in the distribution of profits (for firms with one business with

[15] This additional condition is required to eliminate one of the parameters in the general solution to the differential equation. It says that in the limit, an increase in ω_i will still increase v, while the slope is less than ∞. This is intuitive: When ω_i gets very high, the process is very far away from the optimal stopping point ω^* (which is some negative number). A change $\Delta\omega_i$ in ω_i will then change the value of the project in expected terms approximately by $\frac{1}{\rho}\Delta\omega_i$, i.e. by the annuity value of the present change. So $\lim_{\omega_i \to \infty} v'(\omega_i) < \infty$.

[16] The exit threshold ω^* is also increasing in ρ and decreasing in μ. The less patient the firm is, the lower the (option) value it attaches to future payoffs making it optimal to exit at a higher threshold value. The higher the drift rate of the process, the higher the expected future payoffs, which increases the option value of waiting and decreases the optimal exit threshold.

Figure 1-14: Optimal stopping point when the profit flow follows a simple Brownian motion



constant size) gets longer as the variance in the stochastic evolution of profit flows increases.

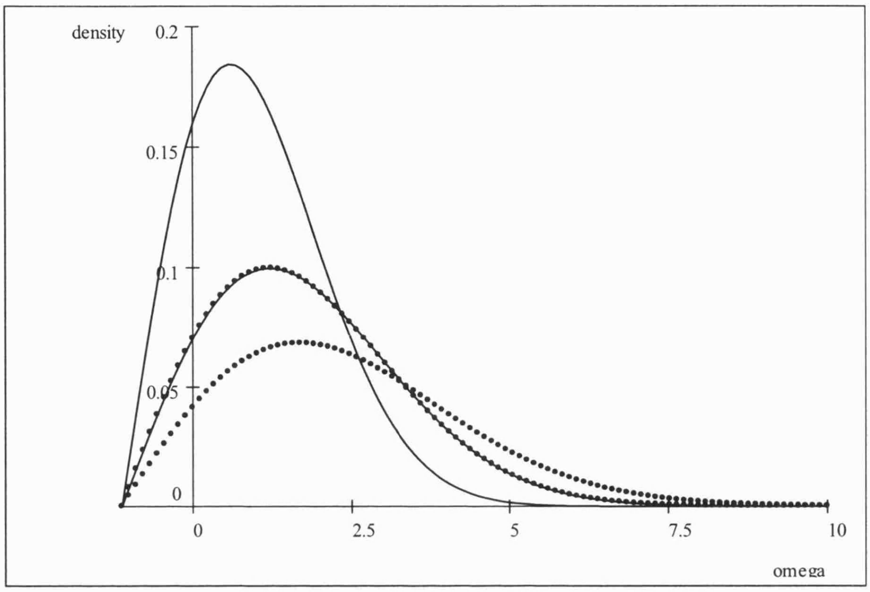
Next, it is useful to examine the distribution of the profit flow of a business over time conditional on survival. To do so, note that the optimal stopping point ω^* acts as an absorbing barrier for the diffusion process of the state variable: if the state hits ω^*, the process is stopped. Figure 1-15 shows the distribution of ω_i conditional on survival for the parameter constellation $(\omega_0 = 0, \mu = 0, \sigma = .5)$ for three points in time $t = \{10, 20, 30\}$. The optimal stopping point at this parameter constellation is $\omega^* = -1.118$. The solid line corresponds to $t = 10$, the dash-dotted line to $t = 20$, and the dotted line to $t = 30$.[17] It is clear from Figure 1-15 that the distribution of the profit flows conditional on survival gets flatter as t increases, even though the underlying process has zero drift.

The following conclusions can be drawn from Figures 1-14 and 1-15: While

[17] When the profit flow follows a Brownian motion with an absorbing barrier in the form of the optimal stopping point at, the distribution of the profit flow conditional on survival can be derived as a mixture of two normal distributions (Cox & Miller 1965, p219-221). With starting point at ω_0 and an absorbing barrier at ω^*, the density of ω at time t is given by

$$f(\omega|t, \omega_0, \omega^*, \mu, \sigma) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi t}}\left(e^{-\frac{(\omega-\omega_0-\mu t)^2}{2\sigma^2 t}} - e^{\frac{2\mu(\omega^*-\omega_0)}{\sigma^2} - \frac{(x-2\omega^*+\omega_0-\mu t)^2}{2\sigma^2 t}}\right) & \text{if} \quad \omega \geq \omega^* \\ 0 & \text{if} \quad \omega < \omega^* \end{cases}$$

Figure 1-15: Distribution of a simple Brownian motion with absorbing barrier over time



there is a negative tail in the distribution (and while this tail gets longer with increases in σ), the "positive tail" is even longer. What is more, over time the positive tail increases in length as there is no mechanism keeping the process from "wandering off towards infinity". Also, the mode and the expectation of the conditional distribution moves up over time. This is a consequence of keeping the process alive if it proves profitable ("preserving the upside risk"), while killing the project if it becomes too unprofitable ("limiting the downside risk"). While these properties are quite natural upon reflection, they are unfortunate as they imply that this simple process will not be able to qualitatively reproduce the shape of the empirical profits-size distributions which suggests that if anything, the negative tail should be longer than the positive tail.[18]

[18] One may think that introducing a negative drift in the process may keep it from warndering off towards infinity. However, a negative drift will also have the effect to shift the optimal stopping point ω^* upwards. Overall, the desired properties could not be generated with this simple specification.

1.3.3 A better alternative: "Costs" following a geometric Brownian motion

To address the shortcomings of the simplest Brownian motion specification, an upper bound on the profit flow is required. To do so consider the following specification which places an upper bound of unity on the profit flow:

$$\pi(\omega_i) = 1 + \omega_i, \text{ where } \omega_i \in (-\infty; 0].$$

One interpretation of this specification is that the business generates a constant revenue stream of unity whereas $-\omega_i$ measures the operating costs of the business.

The stochastic evolution of the profits will then be due to random fluctuations in operating costs. As costs are non-negative and can potentially grow without bounds, it is convenient to formulate their evolution as a geometric Brownian motion, i.e.

$$\frac{d(-\omega_i)}{-\omega_i} = \mu dt + \sigma dz$$

with drift rate μ, variance σ^2. dz denotes the increments of a standard Wiener processes ($E(dz) = 0$, $E(dz^2) = dt$).

This specification of the profit function and the evolution of costs has two immediate consequences. (i) It will impose an upper bound on the profit flow of unity, which is ad hoc, but which can be motivated by interpreting the profit flow as operating margin on sales. (ii) Modelling operating costs as a geometric Brownian motion implies that the variance in the change of the profit flow over a given time horizon increases, the lower the profits get (i.e. the higher the costs). This seems to be in line with the empirical observation about year on year profit rates (see Section 1.2.3 and Figure 1-13).

In the continuation region, the Bellman equation (1.2) now becomes[19]:

$$v(\omega_i) = (1 + \omega_i)dt + \frac{1}{1 + \rho dt} E[v(\omega_i + d\omega_i)|\omega_i]$$

Using Ito's Lemma, rearranging, dividing by dt, and letting $dt \to 0$ now yields

[19] Details of the derivation of this model are given in Appendix A1.3.

the second order partial differential equation:

$$\rho v + \mu \omega_i v' - \frac{1}{2}\sigma^2 \omega_i^2 v'' = 1 + \omega_i \tag{1.6}$$

The value matching and smooth pasting conditions (1.3) still have to be satisfied and, as before, one additional condition is required to eliminate one parameter in the general solution for (1.6). This condition is derived by realising that when ω_i reaches zero, it remains there and the value of the business becomes just the annuity value of of a unity profit flow each period. This implies the additional condition

$$v(0) = \frac{1}{\rho}.$$

Solving the differential equation (1.6) subject to these three conditions yields the value function $v(\omega_i)$ and the optimal stopping point for the state variable ω^*:

$$\begin{aligned} v(\omega_i) &= -\frac{\omega^{*-r_1}}{\rho(r_1 - 1)}\omega_i^{r_1} - \frac{1}{\mu - \rho}\omega_i + \frac{1}{\rho} && (1.7) \\ \text{with } r_1 &= \frac{1}{2\sigma^2}\left(\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}\right) > 0, \text{ and} \\ \omega^* &= -\frac{\rho - \mu}{\rho}\frac{\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}}{-\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}} && (1.8) \end{aligned}$$

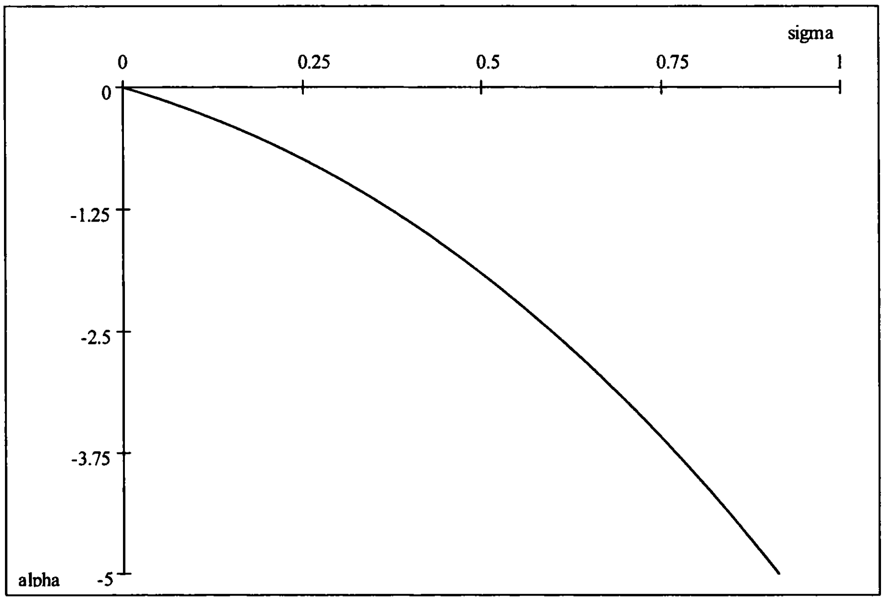
When the state variable of the business ω_i reaches the critical level ω^*, the firm optimally decides to abandon the business, i.e. is forgoes the option of waiting for potential future improvements in the profit flow. Note that this translates into an optimal stopping point in terms of the current profitability of the business as:

$$\alpha(\mu, \sigma, \rho) = 1 + \omega^* = 1 - \frac{\rho - \mu}{\rho}\frac{\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}}{-\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}}. \tag{1.9}$$

Figure 1-16 shows the optimal stopping point for the profit level $\alpha(\mu, \sigma, \rho)$ as a function of σ for $\mu = 0$ and $\rho = .1$. The optimal stopping point clearly decreases as σ increases. The intuition is the same as before: The higher the variance of the process, the higher the upside risk for future profits, i.e. the higher is the option value of delaying the shut down decision.

The distribution of the profits of a business given its initial state ω_0 now evolves

Figure 1-16: Optimal stopping for the geometric Brownian motion model



as a diffusion process with an absorbing barrier at α. Another insight from the simple Brownian motion model in the previous subsection carries over. Even when the expectation of a change in costs of the "unconstrained" process is zero, the expected profit flow of a business conditional on survival increases – even if it is bounded from above by unity with this specification.

1.3.4 Simulations

The analytic solution of the previous section for the optimal stopping point of the profit flow α, can now be used to simulate the stochastic evolution of firms in our model along the two dimensions of profitability and size by allowing for both, random arrival and optimal abandonment of businesses. We model the arrival of new business opportunities as a Poisson process with arrival rate λ. The simulations will combine the option value mechanism with respect to the shut down decision of a single business with the diversification effect arising from averaging across constituent businesses. This will then serve to generate profit-size distributions reflecting the interactions of these two effects. The primary interest here is on whether varying the variance parameter σ, will be sufficient to reproduce the qualitative differences between high and low R&D industries with respect to the profits-size distributions and the intra distribution dynamics.

Firm size in the simulations will be measured by the number of constituent businesses of a firm. Firm profitability is simply the average profit flow across the constituent businesses. This implicitly assumes that all constituent businesses of a firm are of the same size – regardless of whether size represents asset values or sales revenue. In the asset interpretation businesses would generate different profits from sets of assets of the same value. When size is interpreted as sales revenue, the assumption is that each business generates the same revenue but that the operating margin on this revenue differs across businesses. While this assumption is clearly violated in practice, it is the simplest possible specification. The "fit" of this size measure with the empirical data will be discussed below.

To run the simulations, it is necessary to approximate the continuous time process by a discrete time process. We chose a time increment of $1/12$ so that each yearly observation in the simulation consists of 12 discrete steps. (The managers of the firms decide on a monthly basis which businesses projects they want to shut down). Two simulations are run, one with a high variance of the underlying process and one with a low variance. All parameters are to be understood as yearly values. However, since the focus here is on seeing whether the model can qualitatively generate the features of the profits-size distribution and of the intra-distribution dynamics, no attention is paid whether these values are plausible empirically (beyond generating the desired distributions). Apart from the variance parameter σ, all parameters are held constant across the two simulations. After a few experiments the parameters were set to the following values:

Arrival rate of new business:	$\lambda = .2$
Initial profit flow of new business:	$\pi_0 = 0$, i.e. $\omega_0 = -1$
Discount rate:	$\rho = .05$
Drift rate for geometric BM:	$\mu = 0$
Time increment for simulation:	$dt = 1/12$
Number of firms :	$n = 1500$
Maximum age of firms in years:	$m = 50$

The variance parameters were chosen so that the locations of the optimal stopping points were in the correct order of magnitude:

	High Variance	Low Variance
Variance parameter σ	.3	.15
Optimal stopping point for profit flow α	-1.5	$-.6$

After each simulation, the monthly observations for size and profitability are averaged into yearly values. As a result, non-integer values for size arise if the number of businesses of a firm changes within the year.

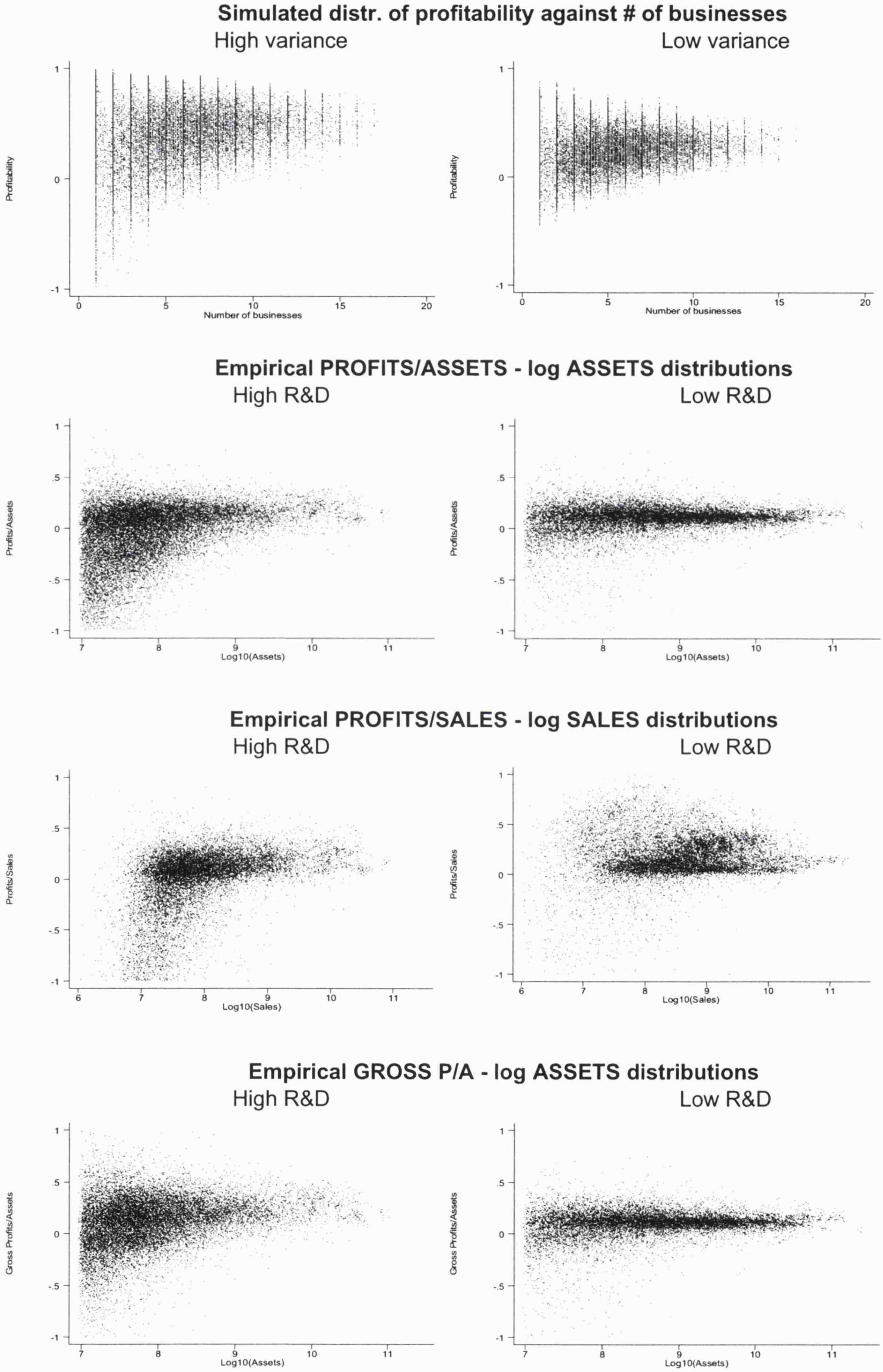
To generate a population of firms of different ages we draw a year s for each firm from a uniform distribution from 1 to 50. In the results presented below on the cross sectional distribution, only observations on simulated firms with $t \geq s$ are included. The transition matrices in the next subsection are based on the years s and $s+4$, i.e. a four year horizon as in the empirical transition matrices

Simulation results 1: The cross sectional distribution

Figure 1-17 shows the scatter plots resulting from the two simulations (top two panels), and the empirical scatter plots for profits/assets (second row), profits/sales (third row), and gross profits/assets (bottom row) as a reference. The left hand side panels correspond to the high variance simulation (respectively the high R&D subsample) and the right hand side panels to the low variance simulation (respectively the low R&D subsample).

An important remark is in order at this stage. While the size dimension in the empirical scatter plots is based on the logarithm of assets (respectively log sales) which implies that movement along this axis represents growth in assets/sales over orders of magnitude, the size dimension in the simulations is a mere count of the number of businesses. That is, the scale is logarithmic in the empirical scatter plots, whereas it is linear in the number of businesses in the simulated plots. This is clearly a major point of disagreement between the simulated and the empirical distributions. We believe that this is partly due to the simplicity of the theoretical model, which might be overcome by alternative specifications such as allowing for growth of individual businesses. While not within our modelling specification, it is plausible that, on average, bigger firms not only consist of more businesses, but also of bigger businesses. (Microsoft consists surely of more businesses than a

Figure 1-17: Simulated and empirical profits-size distributions



small software company, but its businesses (Windows, Office, Xbox, etc.) certainly also operate at a scale which is on average orders of magnitude larger than the businesses of small rivals. As our simple model does not account for the fact that individual businesses differ in terms of size, we abstract from this feature.

However, apart from the different scaling on the size axis, the simulated scatter plots reproduce the empirical scatter plots remarkably well. The strong negative tail in the high variance simulation is strongly pronounced for firms with few businesses and disappears quickly as the number of businesses increases. This mimics the shape of the empirical negative tail in the high R&D subsample. In comparison, the negative tail in the low variance simulation is significantly shorter (as implied by the endogenously higher value of the optimal stopping point α) and the overall dispersion is lower. This shortening of the negative tail and the decrease in the dispersion as one moves from the high variance (Figure 1-17, top left panel) to the low variance simulation (top right panel) reproduces quite well the qualitative difference between the profits/assets - log assets scatter plots (second row). The low R&D scatter plot for profits/sales - log sales (third row, right panel) looks puzzling as there are surprisingly many dots with very high profits over sales. A look at Figure 1-6 suggests, that this is due to the high dispersion of profits/sales in the "Crude Petrol and Natural Gas" industry (SIC 1311). Finally, the simulated scatter plots also look also relative similar to the empirical scatter plots for gross profits/assets (bottom row).

To investigate the fit of the simulated distribution with the empirical distributions a little further, Figures 1-18 and 1-19 present histograms for profits for small and big firms across all the settings. The top panels in each figure show histograms on simulated profitability, the second row panels on profits/assets, the third row panels on profits/sales, and the last row for gross profits/assets. The left hand panels pool over all small firms and the right hand panels over large firms. "Small" is defined as having up to 6 businesses for the simulations and below 1 billion US$ in assets/sales for the empirical graphs.

Figure 1-18 shows the high R&D/variance settings and Figures 1-19 the low R&D/variance settings. These pictures show that, while the qualitative differences in terms of dispersion between small and large firms and low and high variance firms

Figure 1-18: Histograms on profitability for HIGH variance/R&D group

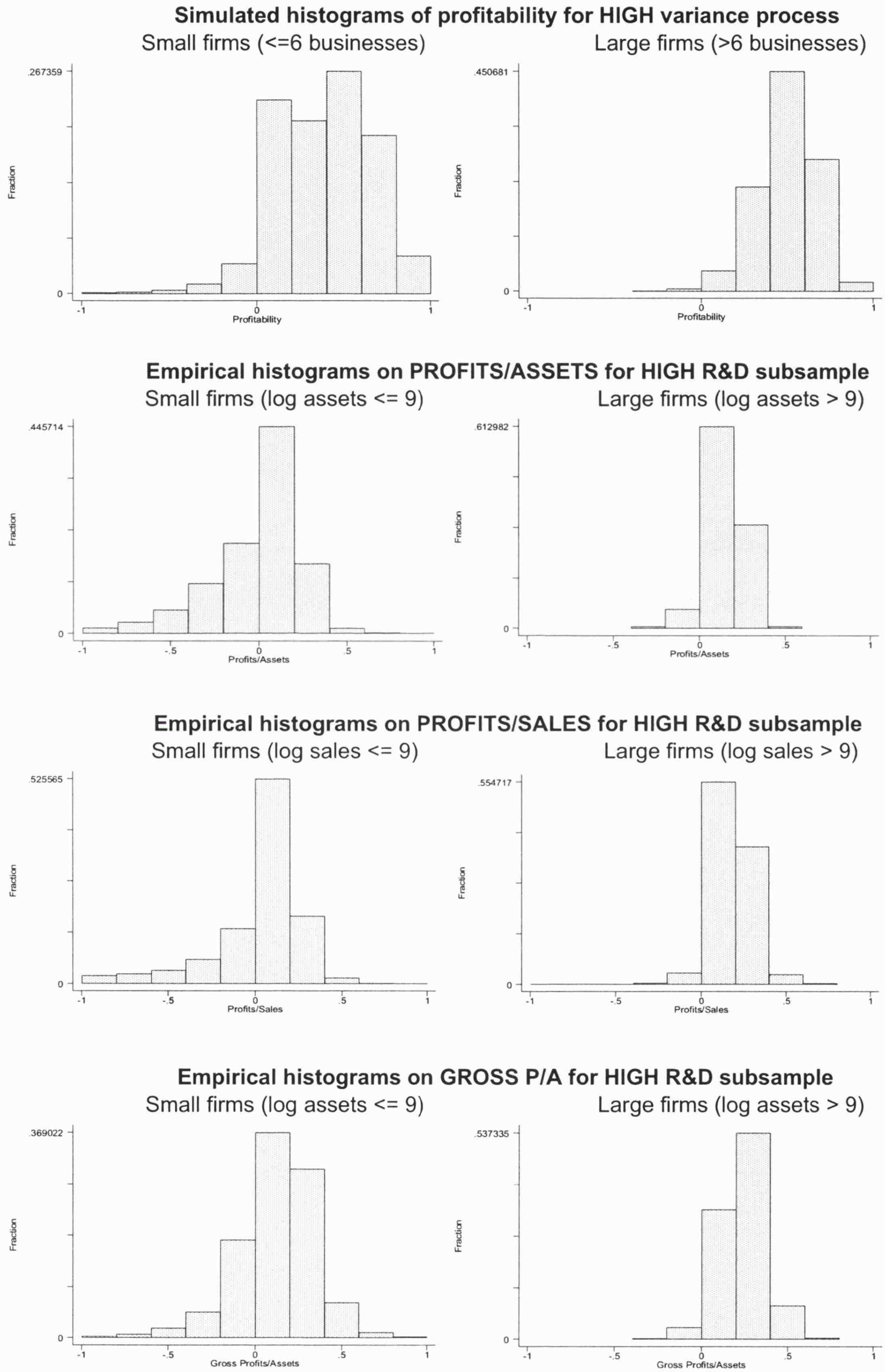
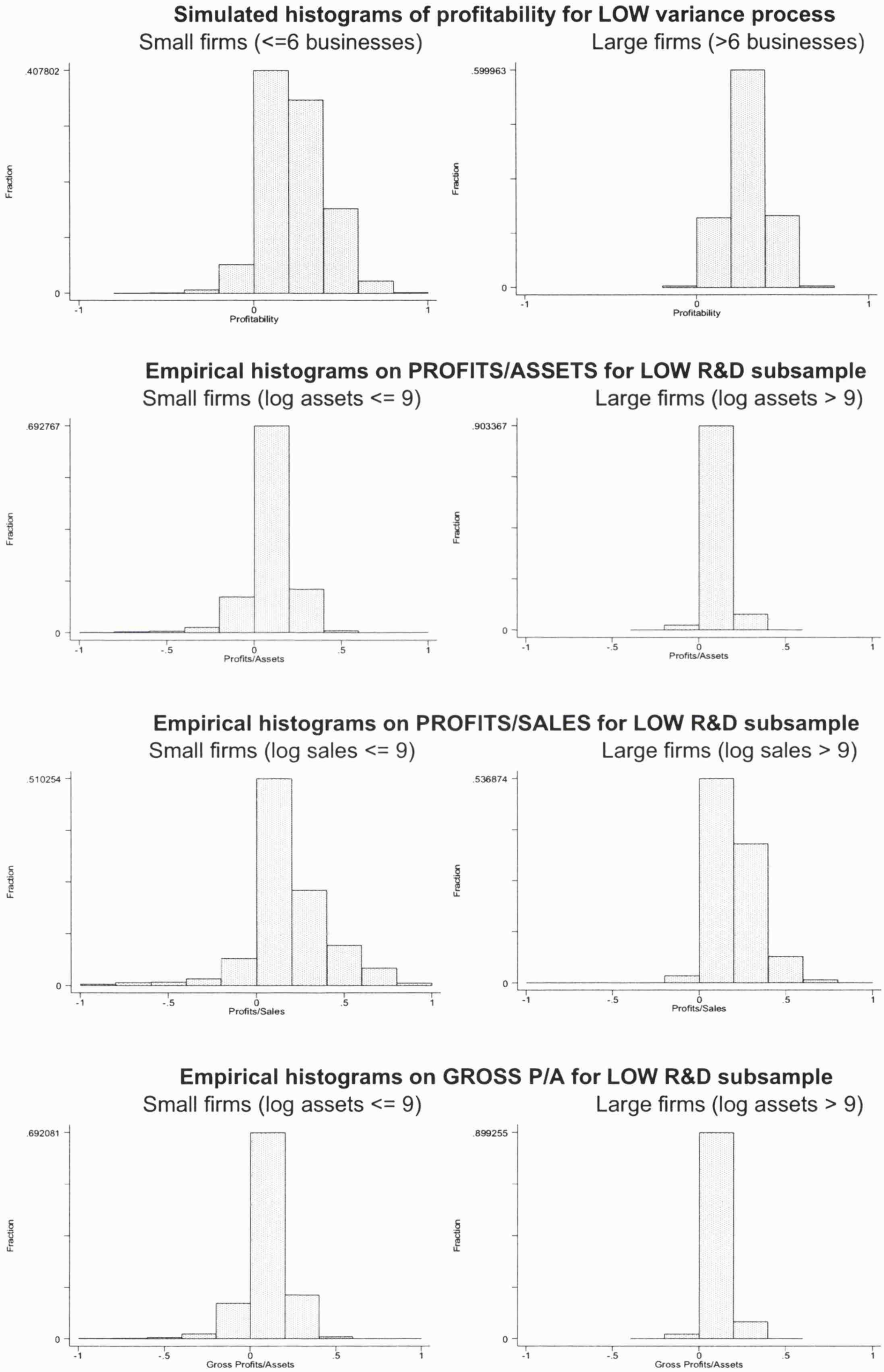


Figure 1-19: Histograms on profitability for LOW variance/R&D group



are roughly generated by the model, the mode of the histograms is higher in the simulations than in the empirical histograms. The simulated profit rates of the median firms are too high. This feature might be due to the property of the diffusion process with an absorbing barrier that, conditional on survival, the mode of the profits distribution of a single business "shifts up" over time.

Why is one year equal to one period in the model

Simulation results 2: Stylised intra distribution dynamics

How is it calib

Would result differ if t=8 rather than t=4

The simulated paths of firms can also be used to compare the differences in the intra distribution dynamics across the two simulations with the stylised qualitative differences between the distribution dynamics for high and low R&D industries discussed above. Tables 1.11 and 1.12 present transition probability matrices for the high and the low variance simulations over 3 active states and one exit state. As with the empirical transitions, two different definitions for the active states are examined – an absolute definition and a relative definition. In both cases, "large firms" are defined as having more than 6 businesses. The absolute definition further distinguishes between small profitable and small unprofitable firms. The relative definition divides the set of small firms in firms with profits over assets above the median small firm in the simulation and firms below the median firm, i.e. it uses the position relative to the median small firm.

"Absolute" states: The striking feature in the empirical transition matrices based on absolute states is that small, unprofitable firms in high R&D industries are more likely to remain in this state over the given time horizon, and less likely to exit and to become profitable than firms in low R&D industries. The model fails to reproduce this feature (Table 1.11):

- High variance small unprofitable firms are significantly more likely to become profitable than low variance firms, why?
- they are significantly less likely to ~~become~~ remain unprofitable, and ✓
- there is little difference with respect to exit rates

The first point is the only feature that is in accordance with the empirical transition matrices and which only becomes significant in the empirical transition matrices once the Pharmaceutical industries have been excluded.

"Relative" states: The qualitative differences in the simulated transition probabilities between high and low variance firms with "relative" states are also somewhat at odds with the empirical differences across the two subsamples (Table 1.12):

Table 1.11: Simulated transition probability matrices for ABSOLUTE states

(Transition probabilities with SE's below - both in %)

Panel a: LOW variance simulation

ABSOLUTE STATE in t		ABSOLUTE STATE IN t+4: large	small prof.	small unprof.	exit	#firms
large	**TP**	**92.8**	**7.2**	**0.0**	**0.0**	**208**
	SE	1.8	1.8	0.0	0.0	
small prof.	**TP**	**10.6**	**83.9**	**5.5**	**0.0**	**886**
	SE	1.0	1.2	0.8	0.0	
small unprof.	**TP**	**0.7**	**50.3**	**43.0**	**6.0**	**149**
	SE	0.7	4.1	4.1	2.0	
Total		**23.2**	**67.0**	**9.1**	**0.7**	**1243**

Panel b: HIGH variance simulation

ABSOLUTE STATE in t		ABSOLUTE STATE IN t+4: large	small prof.	small unprof.	exit	#firms
large	**TP**	**95.5**	**4.5**	**0.0**	**0.0**	**314**
	SE	1.2	1.2	0.0	0.0	
small prof.	**TP**	**11.3**	**82.9**	**5.6**	**0.2**	**852**
	SE	1.1	1.3	0.8	0.2	
small unprof.	**TP**	**2.9**	**64.8**	**25.2**	**7.2**	**139**
	SE	1.4	4.1	3.7	2.2	
Total		**30.7**	**62.1**	**6.4**	**0.9**	**1305**

large: # businesses > 6
small, profitable: profits>0 & # businesses <= 6
small, unprofitable: profits<=0 & # businesses <= 6

Table 1.12: Simulated transition probability matrices for RELATIVE states

(Transition probabilities with SE's below - both in %)

Panel a: LOW variance simulation

RELATIVE STATE in t		RELATIVE STATE in t+4: large	small hi prof.	small lo prof.	exit	#firms
large	**TP**	**92.8**	**6.3**	**1.0**	**0.0**	**208**
	SE	1.8	1.7	0.7	0.0	
small hi prof.	**TP**	**14.2**	**66.5**	**19.3**	**0.0**	**471**
	SE	1.6	2.2	1.8	0.0	
small lo prof.	**TP**	**5.0**	**24.7**	**68.8**	**1.6**	**564**
	SE	0.9	1.8	2.0	0.5	
Total		**23.2**	**37.4**	**38.7**	**0.7**	**1243**

Panel b: HIGH variance simulation

RELATIVE STATE in t		RELATIVE STATE in t+4: large	small hi prof.	small lo prof.	exit	#firms
large	**TP**	**95.5**	**2.9**	**1.6**	**0.0**	**314**
	SE	1.2	0.9	0.7	0.0	
small hi prof.	**TP**	**14.5**	**63.9**	**21.7**	**0.0**	**415**
	SE	1.7	2.4	2.0	0.0	
small lo prof.	**TP**	**6.9**	**28.5**	**62.5**	**2.1**	**576**
	SE	1.1	1.9	2.0	0.6	
Total		**30.7**	**33.6**	**34.9**	**0.9**	**1305**

large: # businesses > 6
small, high profit: profits > median small firm & # businesses <= 6
small, low profit.: profits <= median small firm & # businesses <= 6

- The only quantitatively significant difference in the two simulated transition matrices is that high variance, low profits firms are less likely to stay in this state and more likely to become high profits firms. This is, if anything, the opposite of what emerges from the empirical transition matrices.

Summary

Overall, the simple model with independent businesses of constant size and independent shut down decisions for each business captures the key features in the cross sectional profits-size distributions for high and low R&D industries. There are, of course, a number of points where the model is at odds with the data: While reproducing the "negative tail" and its disappearance and the differences in the profits-size distribution between low and high R&D (by means of varying a single parameter for the variance of the stochastic process driving the profit flow for a business), the mode of the implied profit rates is generally higher than that of the empirical distributions. Moreover and more importantly, the scale along the size dimension differs in the simulated distributions and the empirical distributions. With respect to the intra distribution dynamics, the simulations also show that the model fails to reproduce most "stylised facts" on empirical intra distribution dynamics.

However, given the simplicity of the model which combines only two basic effects of optimal stopping and diversification, the shortcomings may not be surprising and the fact that it generates the qualitative differences in the profits size distribution quite well is an indication that the option value and the diversification effect are the major mechanisms driving this empirical phenomenon.

1.4 Conclusion

This study documents a striking regularity in the cross sectional distribution of firm profitability and firm size: In a group of 18 high R&D industries there is a long negative tail of small loss-making firms in the cross sectional distribution, whereas this tail is much weaker in a control group of 20 low R&D industries. Moreover, the cross sectional variance in profit rates declines with firm size in both groups of industries. We believe that these distributional features are due to a combination of two effects: A real option effect explaining existence of a large number of small loss-making firms in high R&D industries; and a diversification effect explaining the decline in the variance of profit rates with firm size.

We propose a simple stochastic reduced form model for firm dynamics that combines these two effects and show that this model can generate the qualitative features of the profits-size distribution and the variation across high and low R&D industries by varying the variance in the underlying stochastic process. However, the scale of the size dimension in the model does not agree with that in the data. Unfortunately, the model also fails to reproduce the most interesting stylized facts about intra-distribution dynamics. In particular, it does not generate the empirical observation that small unprofitable firms in high R&D industries are more likely to remain unprofitable than firms in low R&D industries.

Given the simplicity of the model and its reduced form nature, this is perhaps not surprising. Industry dynamics are certainly more complex than the simple model in this chapter which, because of its reduced from nature, abstracts from many firm decisions such as investments in physical capital and R&D. Modelling these firm decisions and estimating underlying stochastic processes driving them is a challenging task to which we will turn in the next chapter. However, we believe that this simple model proves that the observed qualitative differences in the profits-size distributions are due to differences in the riskiness of the economic environment and to a combination of a real option and a diversification effect.

Chapter 2

R&D and the Dynamics of Productivity

2.1 Introduction

This chapter provides an empirical framework for the analysis of investments in research and development (R&D) on the entire distribution of a firm's future productivity conditional on its current productivity. The approach builds on a structural model for firm dynamics in which firms invest in R&D and in capital. We show that the investment policy function generated by this model is invertible. This allows us to use a production function approach similar to Olley & Pakes (1996) to estimate the firm's unobserved productivity states. The main empirical focus lies on using these estimates to determine the evolution of productivity conditional on current productivity and R&D.

The conditional distribution of future productivity is at the heart of understanding the dynamics of firms in research and development (R&D) intensive industries and the incentives to invest in R&D as well as in physical capital. In theoretical models of firm dynamics (e.g. Ericson & Pakes (1995), Hopenhayn (1992)), the success or failure of a firm in an industry and its dynamics along other characteristics is typically driven by the stochastic evolution of a firm specific productivity state which the firm may or may not be able to influence through R&D investments. In these models, the distribution of a firm's future productivity state conditional on its current productivity state and its R&D investments is therefore the key stochastic primitive driving firm dynamics, investment incentives, and idiosyncratic differences in outcome paths across firm. The framework in this chapter allows us to estimate an empirical counterpart of this central primitive of the theoretical literature of firm dynamics.

The question of the effect of R&D on productivity is, of course, the theme of a much older and huge empirical literature initiated by Zvi Griliches (for an excellent review see e.g. Griliches (1998)). This literature has largely been concerned with estimating the average or expected (private or social) returns to (firm or industry level) R&D. While estimating the average effect of R&D on productivity is often the best one can do, an analysis of the effect of R&D on the entire distribution of future productivity at the firm level clearly provides a more complete picture and makes explicit the stochastic nature of the outcomes of R&D. Investigating this distribution provides explicit information on the stochastic environment in the industry under study and therefore provides a better understanding of the private incentives to invest in R&D. Although the literatures on firm dynamics and on R&D and productivity are closely related from this perspective, this link has received relatively little explicit attention.

value add?

In their seminal paper on the Telecommunications Equipment industry, Steven Olley and Ariel Pakes (1996) propose a new semiparametric method for the estimation of production functions in the presence of unobserved, serially correlated productivity differences across firms. Their approach overcomes well known biases in the parameter estimates of the production function which are due to the simultaneity of input choices and selection through exit.

The approach of Olley and Pakes relies on a structural single agent model for the dynamics of firms in the industry. In their model, firms maximise their expected discounted value by deciding on whether to exit the industry and on how much to invest in physical capital. The state variables of the firm are the capital stock and an unobserved productivity state. While the firm can control the capital stock through its investment decisions, the evolution of the unobserved productivity state is driven by an exogenous first order Markov process. The key feature of their model (proven in Pakes (1994)) is that it generates an investment policy function that is increasing in the unobserved productivity conditional on the firms capital stock and which is hence invertible.

The invertibility of the policy function allows them to control for the unobserved firm productivity nonparametrically without having to solve the structural model explicitly. Olley & Pakes (1996) provide an estimation algorithm that yields consistent estimates of the production function parameters (which are a subset of the parameters of the underlying structural model) and, as a by-product, also produces estimates of the firm specific unobserved productivity state over time. They use their estimates to analyse the evolution of the aggregate industry productivity over time and find that productivity increases were primarily a result of a reallocation of capital towards more productive establishments.

The present study extends this approach by explicitly allowing for R&D investments by firms. This is an important step: In many industries (including Telecommunications Equipment) firms engage in regular and often heavy R&D investments with the aim of improving future productivity. Without an explicit model, it is unclear whether the Olley-Pakes approach can be applied if one believes that the true underlying model for firm dynamics should include R&D investments.

We make three main contributions: First, we add an R&D investment decision to the controls of the theoretical model and let the stochastic evolution of the unobserved productivity state be influenced by R&D investments as well as by the firm's current productivity state – i.e. we partly endogenise the Markov process for the productivity dynamics. We show that, under certain restrictions on the model primitives, the policy function for capital investments generated by the extended underlying structural model is still invertible. This is a powerful result as it implies that the unobserved productivity state can still be expressed as a function of capital and investment even if the model includes R&D. This is crucial for controlling for the unobservable state variable in empirical work.

Second, the invertibility result suggests that the estimation approach of Olley & Pakes (1996) for the estimation of production function coefficients can be applied even if the true underlying model includes investments in physical capital as well as in R&D. This follows immediately for the first stage of the Olley-Pakes approach which estimates the production function coefficients for variable factors of production. However, the second stage of the estimation procedure which yields coefficient estimates for the quasi-fixed factors requires some further analysis. This is because of the need to control for the expectation of productivity conditional on past information and on survival to consistently estimate the coefficients of the quasi-fixed factors. In a model with R&D, this expectation not only depends on past productivity, but also on past R&D investments. Ignoring the impact of R&D may therefore lead to inconsistent coefficient estimates for the quasi-fixed factors. We present two alternative modifications of stage two of the Olley-Pakes algorithm that address this problem. The first suggests using R&D data but is problematic if the level of R&D endogenous in the sense that it is correlated with potential measurement error in the dependent variable. The second approach does not require any additional data and exploits the fact that we can control for the effect of R&D by estimating stage two as a fully nonlinear nonparametric function of past productivity and current capital. A further analysis of our second modification shows that it is asymptotically equivalent to the second stage estimation equation originally proposed by Olley & Pakes. This is because, in the context of our model, the survival probability can be shown to be strictly increasing in capital conditional on productivity. As a result, both stages of the original Olley-Pakes procedure can be directly applied to our model with R&D.

While this equivalence result is specific to the setup of our structural model, the problem of controlling for R&D is more general: Whenever the expected future productivity depends on R&D (i.e. R&D has some effect), one needs to think carefully about how to control for R&D in stage two of the estimation procedure to obtain consistent estimates of the quasi-fixed factors. This is independent of whether one uses investment to proxy for productivity differences as Olley & Pakes (1996) or whether one employs alternative proxies such as intermediate inputs as proposed by Levinsohn & Petrin (2003). The intermediate input approach in Levinsohn & Petrin (2003) is attractive because it does not require a monotonicity property derived from a structural dynamic model to derive a proxy for unobserved productivity. However, if the underlying model includes R&D, one still has to write down the candidate structural model to explore how to adequately correct the estimation procedure to yield consistent estimates for the quasi-fixed inputs.

Our third, and main empirical, contribution is that we use the firm and year specific productivity estimates from the production function estimation to analyse the distribution of future productivity conditional on current productivity and R&D. We can test whether the estimated conditional distribution satisfies the first order stochastic dominance assumptions of the theoretical model (which are standard in the literature on firm dynamics) and hence whether the model is rejected by the data. By providing a method of examining the effect of R&D on the entire distribution of future productivity and by quantifying this effect, we also hope to contribute to the empirical literatures on firm dynamics and on R&D and productivity.

We use firm level COMPUSTAT data for the four 3-digit SIC industries "Pharmaceuticals (SIC 283)", "Computer Hardware (SIC 357)", "Telecommunications Equipment (SIC 366)", and "Software (SIC 737)". The sample spans the years 1980 to 2001 and is characterised by high levels of R&D spending and by a considerable degree of exit. In fact, the mean and median levels of R&D investments exceed the corresponding levels of investment in physical capital in each of these industries. This suggests that controlling for R&D as well as survival is potentially important. Estimates for the capital coefficient are sensitive to the specification of the stage two estimation equation. In particular, the point estimates for capital from our fully nonlinear specification are lower than those estimated from the original Olley-Pakes equation in all industries except "Software". While these differences are not statistically significant, it indicates that despite the asymptotic equivalence of the approaches their finite sample performance may differ.

Specification tests proposed in Olley & Pakes (1996) lead us to accept the model for "Pharmaceuticals" and "Telecom Equipment" but to reject it for "Computers" and "Software". Testing whether the future productivity distributions conditional on R&D and current productivity satisfy the first order stochastic dominance property of the model leads to the same conclusion. For the industries "Computers" and "Software", this suggests that using investment as a proxy for productivity does not adequately control for unobserved productivity differences across firms and that the structural dynamic model does not adequately capture the dynamic features of these industries. However, the estimation approach and the model seem to work well for "Pharmaceuticals" and "Telecom Equipment".

Further analysis of the future productivity distribution conditional on R&D and current productivity shows that productivity is more volatile in "Telecom Equipment" than in "Pharmaceuticals" as shown by a higher dispersion of productivity increments. We estimate the average elasticity of next period's productivity with respect to R&D to be around .02 for "Pharmaceuticals" and, depending on the specification between .007 and .04 for "Telecom Equipment". These effects are low but significant and represent estimates of the returns to R&D in terms of productivity from one year to the next.

The chapter is organised as follows. Section 2.2 presents the model, assumptions, and theoretical results that form the basis of the empirical approach. Section 2.3 then discusses the Olley-Pakes estimation approach and proposes modifications of it that account for the model with R&D. Section 2.4 introduces the data. In section 2.5, we present and discuss production function estimates, whereas section 2.6 empirically investigates the conditional productivity distribution. Section 2.7 concludes. The appendix contains proofs and details on data construction.

2.2 Theoretical framework

2.2.1 Structure of the model

We use a stochastic dynamic single agent model for the industry and assume that firms maximise the expected discounted value of future net cash flows. Firms have two state variables: productivity ω and the capital stock k. At the beginning of each period, each firm observes its state and makes a discrete decision whether to exit or continue in operation. If it exits, it receives a termination value Φ. | const Otherwise, it earns current period profits $\pi(\omega, k)$ and decides how much to invest in physical capital and in R&D.

Capital investment has a deterministic effect on the future capital stock, while R&D influences future productivity stochastically. Instead of formulating the firm's decisions in terms of capital investment and R&D expenditure, we will setup the model such that the firm directly chooses the next period's capital stock (through its capital investments) and the distribution of next period's productivity (through its R&D spending).

Capital k accumulates deterministically according to the usual equation

$$k' = (1 - \delta)k + i,$$

where k' denotes next period's capital stock, δ the rate of capital depreciation and i the investment choice of the firm. In choosing the amount of investment i, the firm therefore effectively chooses the next periods capital stock k'. The cost of physical capital investment of achieving this k' depends on the capital state of the firm and is denoted $c(k', k)$.

The productivity state ω evolves stochastically over time according to a controlled Markov process, where the distribution of next period's productivity is increasing (in the first order stochastic dominance sense) in the current productivity state ω and in the amount of R&D expenditure: Conditional on R&D, higher current productivity states result in better future productivity distributions so that there is a degree of persistence in productivity over time. Similarly, conditional on current productivity, the distribution of next period's productivity will be increasing in the amount of R&D investments.

Rather than modelling this idea as a choice of R&D investments, we let the firm choose the distribution of its future productivity from a menu of distributions. The distribution of next period's productivity ω' is a member of the family of distributions

$$\mathcal{F}_{\psi'} = \{F(\omega'|\psi'), \psi' \in \Psi\}$$

which are stochastically increasing in ψ' in the first order stochastic dominance sense. In each period, the firm chooses a distribution ψ' for next period's productivity from this family. The choice of distribution ψ' will require R&D investments of $r(\psi', \omega)$, which are increasing in ψ' and decreasing in ω. So in each period, the firm "buys" its desired distribution for next period through its R&D expenditure, where better distributions come at a higher price but the price of a given distribution is decreasing in current productivity. In this way, we capture the idea that conditional on R&D the future distribution is increasing in productivity ω, and that conditional on ω the distribution is increasing in the amount of R&D.

By introducing a single index ψ' for the distribution F rather than writing the distribution as $F(\omega'|\omega,\text{R\&D})$, we have imposed an important restriction which implies that productivity ω and R&D both affect the distribution for ω' only through ψ'. The members of the family of distributions F are completely ordered (in the first order stochastic dominance sense) by ψ'. This excludes the possibility that R&D and productivity affect F in qualitatively different ways which could lead to a crossing of two distribution functions. A consequence of this restriction is that R&D and ω can be traded off against each other (at least at the margin) in the sense that two firms with different productivity states can have exactly the same distribution for ω' (i.e. the same ψ'), provided the firm with the lower productivity sufficiently outspends the high productivity firm on R&D. Although severe, the theoretical payoff of this restriction is substantial and will become apparent below.

With these specifications in hand, we can formulate the Bellman equation of the dynamic model with discount factor β:

$$V(\omega, k) = \max\left\{\Phi, \sup_{k',\psi'}\left[\pi(\omega, k) - c(k', k) - r(\psi', \omega) + \beta\int V(\omega', k')dF(\omega'|\psi')\right]\right\}. \tag{2.1}$$

At every point in time, the value function $V(\omega, k)$ is a function of the firm's current state vector. The firm's controls are a discrete exit decision, and continuous choices of next period's capital stock k' and next period's productivity distribution ψ'. The maximum expected discounted value of a firm, $V(\omega, k)$, is the larger of two values, the sell-off value Φ and the best possible expected discounted value of continuation. Conditional on choosing the capital stock k' and the next period's distribution ψ' optimally, the continuation value consists of current period profits $\pi(\omega, k)$, reduced by the cost of investment $c(k', k)$ and by R&D expenditure $r(\psi', \omega)$ plus the expected discounted value from the next period onwards.

The solution of the model will yield policy functions for the discrete exit decision and for the continuous capital and distribution choices. The latter can easily be translated into the "conventional" policy functions for physical capital investment and R&D:

$$\text{Exit rule} \ : \ \chi = \begin{cases} 1 & \text{(continue) if } \omega \geq \underline{\omega}(k) \\ 0 & \text{(exit) otherwise} \end{cases} \tag{2.2}$$

$$\text{Capital choice} \ : \ k' = \widetilde{k}(\omega, k) \tag{2.3}$$

$$\text{Distribution choice} \ : \ \psi' = \widetilde{\psi}(\omega, k) \tag{2.4}$$

The particular form of the exit rule in equation (2.2) results from the fact that the profit function is increasing and the R&D function decreasing in ω – assumptions which we will formally introduce in the next subsection. They imply that the continuation value must be increasing in current productivity ω. Since the termination value Φ is constant, the exit rule has the form of a simple exit threshold: For each k, there exists an exit threshold productivity $\underline{\omega}(k)$. If the productivity realisation is below $\underline{\omega}(k)$ the firm exits, otherwise it stays in operation.

Note that this model with R&D and capital investments nests the Olley-Pakes model without R&D. If the choice of distribution is replaced by the current productivity state (i.e. $\psi' = \widetilde{\psi}(\omega, k) \equiv \omega$) and the R&D cost function is set to zero (i.e. $r(\psi', \omega) \equiv 0$), one obtains the Olley-Pakes model.[1]

The empirical work in this study relies on the fact that the productivity state ω (which is observed by the firm but unobserved by the econometrician) can be recovered from the observed investment behaviour of the firms. This requires that the policy function for capital investment (2.3) be strictly increasing and hence invertible in productivity on a known subset of the (k', k) space. A formal proof of this property requires additional assumptions and relies on the supermodularity of the value function. The intuition for the monotonicity of the capital choice in productivity is that the marginal value of capital increases in productivity and that firms with higher productivity states are more likely to have better productivity realisations in the future.

[1] E.g. firms could be "forced" to set $\psi' = \omega$, by an R&D function $r(\omega, \omega) \equiv 0$, $r(\psi', \omega) \equiv \infty$ for $\psi' \neq \omega$.

2.2.2 Assumptions on the model primitives

This subsection formally introduces the assumptions needed for the theoretic results in the next subsection. The assumptions are presented separately for each model primitive. However, they can be classified in four types of assumptions that are necessary for different parts of the proofs below: Assumptions on boundedness are labelled (A1). Assumptions on the monotonicity of the primitives with respect to their arguments are labelled (A2). (A3) refers to supermodularity assumptions[2] and (A4) to differentiability. (A0) are general assumptions on the state space and choice sets.

The state space and the choice sets satisfy:

(A0) $(\omega, k) \in \Omega \times K \subseteq \mathbb{R} \times \mathbb{R}_+$, and
$(k', \psi', \chi) \in K \times \Psi \times \{0,1\} \subseteq \mathbb{R}_+ \times \mathbb{R} \times \{0,1\}$.

The single period profit function $\pi(\omega, k)$ is

(A1.a) bounded from above,

(A2.a) increasing in productivity ω and k,

(A3.a) strictly supermodular in (ω, k), and

(A4.a) continuously differentiable.

The cost of physical capital investment $c(k', k)$ is

(A1.b) bounded from below,

(A2.b) either of the following two

i. increasing in k' and decreasing in k, or

[2]A function $f(x,y)$ with real valued arguments x and y is supermodular in (x,y) if it has increasing differences, i.e if for any $x_1 > x_2$ and $y_1 > y_2$,

$$f(x_1, y_1) - f(x_1, y_2) \geq f(x_2, y_1) - f(x_2, y_2),$$

and strictly supermodular if the inequality holds strictly. When $f(x,y)$ is twice continuously differentiable, f is supermodular iff $\frac{\partial^2 f}{\partial x \partial y} \geq 0$ and we shall adopt that strict supermodularity implies $\frac{\partial^2 f}{\partial x \partial y} > 0$. Finally, $f(x,y)$ is submodular if $-f(x,y)$ is supermodular.

ii. convex in k' with $\min_{k'} c(k', k)$ nonincreasing in k and $\arg\min_{k'} c(k', k) \leq k$,

(A3.b) submodular in (k', k) (i.e. $-c(k', k)$ is supermodular), and

(A4.b) continuously differentiable.

The R&D function $r(\psi', \omega)$ is

(A1.c) nonnegative,

(A2.c) decreasing in productivity ω and increasing in ψ',

(A3.c) submodular in (ψ', ω) and strictly submodular on the set $\{(\psi', \omega)|r(\psi', \omega) > 0\}$, and

(A4.c) continuously differentiable on the set $\{(\psi', \omega)|r(\psi', \omega) > 0\}$.

The distribution of future productivity $F(\omega'|\psi')$ is

(A3.d) strictly stochastically increasing in ψ' in the first order stochastic dominance sense

(A4.d) differentiable in ψ', and invertible in ω'.

The discount factor β and the sell-off value Φ satisfy:

(A1.d) $0 < \beta < 1$,

(A1.e) $|\Phi| < \infty$.

The assumption that $\pi(\omega, k)$ is increasing in both its arguments is standard in this literature. It is required to ensure that the value function is supermodular, which is key for the monotonicity proofs. Supermodularity of the profit function implies that the marginal profitability of capital is increasing in productivity, which provides the main intuition for the monotonicity results. While it is easy to find examples of profit functions that satisfy the supermodularity assumptions, it is also easy to find counter examples that do not.

We allow for two cases for the investment cost function in assumption (A2.b). While the first one is the more straightforward, assumption (A2.b.ii) accommodates the case of adjustment costs which imply that the investment cost function has a minimum at some value of future capital stock which lies at or below the current capital stock. The supermodularity assumption implies that the cost of a given increment in the future capital stock from k_1' to k_2' is decreasing in the current capital. These assumptions are satisfied for most empirical investment cost functions such as quadratic adjustment costs.

That the amount of R&D required is increasing in the future distribution and decreasing in productivity is intuitive and has been discussed above. Submodularity of the R&D function implies that an improvement of the future productivity distribution from ψ_1' to ψ_2' is more costly for low productivity firms than it is for high productivity firms. We assume strict submodularity only on the subset where R&D expenditure is positive. This is intended to deal with the fact that a firm may choose not to invest in R&D at all, in which case the firm will chose the best distributions available at zero cost.

An important simplifying assumption is that the R&D function does not depend on capital. This rules out the possibility that firms' R&D expenditure has to be somehow proportional to their size (i.e. capital stock) to achieve a given effect on the future distribution of productivity. ? true ??

many different types of business ⇒ invest earlier !

The distribution of future productivity is stochastically increasing in the single parameter ψ'. The complete ordering of distributions in this family by ψ' excludes the possibility that R&D has a qualitatively different effect on the distribution from the effect of current productivity, as discussed above. A further restriction is that the distribution does not depend on the capital stock and hence the size measure of the firm. As with the R&D function, such a dependency would be desirable but would destroy the proofs.

2.2.3 Properties of the value and policy functions

We now turn to the analysis of the properties of the value and policy functions. To make empirical use of the model, strict monotonicity of the investment policy function (2.3) on a known subset is required. To prove this property, we will use results from the literature of monotone comparative statics and lattice theory (e.g. Topkis (1978), Topkis (1998), Athey (1995)). This literature identifies supermodularity and its derivatives as a key property in generating robust monotonicity results. While supermodularity is in general a sufficient but not a necessary condition for these results (Milgrom & Shannon (1994)) it easier to work with than quasi-supermodularity which is both necessary and sufficient.

The literature on robust comparative statics is mostly focused on finding conditions on the objective functions under which the argument maximum will be monotonic in a parameter of the objective function. In the present context, the aim is to prove that the optimal future capital stock k' is increasing in the "parameter" productivity ω (for fixed k). However, since our problem is a dynamic programme, the value function enters the objective function through the expectation term in (2.1). Properties of the policy function will therefore naturally also depend on the properties of the value function.

Therefore, we proceed as follows. First, we prove certain properties of the value function in Lemma 2.1 that are useful for the monotonicity results. The monotonicity argument for the capital choice in productivity is then broken down into two steps. The first is the monotonicity of the optimal capital choice in the choice of distribution (Lemma 2.2). We then prove the monotonicity of the distribution choice in productivity (Lemma 2.3). Together, these results yield the weak monotonicity of the investment policy function in Theorem 2.4.

Assumptions (A0) to (A3) are sufficient to prove supermodularity of the value function and weak versions of the monotonicity results below. For strict monotonicity results, the assumptions on strict supermodularity and continuous differentiability (A4) are also necessary.

Lemma 2.1 *The **value function** $V(\omega, k)$ of the model in equation (2.1) is*

1. *bounded under (A0),(A1),*
2. *nondecreasing in productivity ω and k under (A0),(A2),*
3. *unique under (A0),(A1), and*
4. *supermodular under (A0) to (A3).*

Proof. *See Appendix A2.1.* ■

The significance of the introduction of ψ' becomes apparent in the proof of supermodularity, as ψ' is key in showing that the expectation term in equation (2.1) is supermodular if $V(\omega', k')$ is supermodular. The introduction of ψ' also has a huge payoff for the monotonicity results, as it allows us to break down the argument into the monotonicity of the capital choice k' in the choice of future distribution ψ' (Lemma 2.2) and the question whether more productive firms will choose better distributions (Lemma 2.3). Lemma 2.1 puts us in a position to attack these questions.

First, we examine properties of the optimal capital choice conditional on the choice of next period's distribution. Note that, conditional on ψ', the optimal choice of k' does not depend on ω, as ω only enters the profit and the R&D function.

Lemma 2.2 *The optimal* ***capital choice conditional*** *on the choice of distribution ψ' and capital k of the model in equation (2.1),*

$$\kappa(\psi', k) = \arg\sup_{k'} \left[-c(k', k) + \beta \int V(\omega', k') dF(\omega'|\psi') \right],$$

1. *is nondecreasing in ψ' under (A0) to (A3), and*

2. *nondecreasing in k under (A0),(A1), (A3.b).*

Proof. *1. Nondecreasing in ψ': For any given k, the optimal choice k' will be nondecreasing in ψ', provided the objective function is supermodular in (k', ψ') (Topkis (1978) , Theorem 6.1). The integral is supermodular in (k', ψ') because $V(\omega', k')$ is supermodular and $F(\omega'|\psi')$ is stochastically increasing in ψ', so that the result follows (see the argument in the proof of Lemma 2.1).*

2. Nondecreasing in k': For ψ' fixed, $\kappa(\psi', k)$ is increasing in k as $-c(k', k)$ is supermodular implying that the objective function is supermodular in (k', k). ■

Part (1) of Lemma 2.2 corresponds to Lemma 3 in Pakes (1994) and nicely illustrates the intuition behind the monotonicity of the policy function in Olley & Pakes (1996): If the marginal value of capital is increasing in productivity, firms with better future productivity distributions will, all else equal, invest more in capital.

In the model in Olley & Pakes (1996) without R&D, having a better distribution for ω', is equivalent to having higher productivity ω today. In the model with R&D, Lemma 2.2 reflects only the optimal choice for an intermediate problem, which helps reducing the complexity of the subsequent analysis. The second step is provided in Lemma 2.3 which says that more productive firms today will also choose better productivity distributions.

Lemma 2.3 *The policy function in equation (2.4) for the* ***choice of distribution***

$$\widetilde{\psi}(\omega,k) = \arg\sup_{\psi'} \left[\pi(\omega,k) - c(\kappa(\psi',k),k) - r(\psi',\omega) + \beta \int V(\omega',\kappa(\psi',k))dF(\omega'|\psi')\right]$$

1. *is nondecreasing in ω, under (A0),(A1), (A3.c), and*
2. *strictly increasing in ω on the set $\{(\omega,k)|r(\widetilde{\psi}(\omega,k),\omega) > 0\}$, under (A0) to (A4).*

Proof. *1. Nondecreasing: Again, by supermodularity of the objective function in (ψ',ω). The relevant term is $-r(\psi',\omega)$ which is supermodular by assumption.*

2. Strictly increasing: See Appendix A2.1. ■

Strict monotonicity obtains only on the subset of the state space where R&D spending is positive and plays a crucial role in the proof of the strict monotonicity result in Theorem 2.4. This theorem is the empirically powerful result, as it implies that the policy function for next period's capital choice can be inverted to yield the unobserved productivity state ω:

Theorem 2.4 *The policy function for the* ***capital choice*** $\widetilde{k}(\omega,k) = \kappa(\widetilde{\psi}(\omega,k),k)$ *is*

1. *nondecreasing in ω under (A0) to (A3), and*
2. *strictly increasing in ω on the set $\{(\omega,k)|\widetilde{k}(\omega,k) > k \wedge r(\widetilde{\psi}(\omega,k),\omega) > 0\}$ under (A0) to (A4).*

Proof. *1. Nondecreasing: The result follows directly from Lemma 2.2 and Lemma 2.3 as* $\widetilde{\psi}(\omega, k)$ *is nondecreasing in* ω *and* $\kappa(\psi', k)$ *is nondecreasing in* ψ.

2. Strictly increasing: See Appendix A2.1. ■

The fact that the strict monotonicity result obtains on a known subset of the state space implies that on this subset the investment policy function can be inverted:

Corollary 2.5 $\widetilde{\omega}(k', k) \equiv \widetilde{k}^{-1}(k', k)$ *exists on the set* $\{(\omega, k)|\widetilde{k}(\omega, k) > k \wedge r(\widetilde{\psi}(\omega, k), \omega) > 0\}$

Proof. *Follows directly from the strict monotonicity result in theorem 2.4.* ■

This corollary says, that the unobserved productivity state ω can be expressed as function of the current and next periods capital stock on the subset of the data with positive investments in R&D and physical capital. The fact that this function only exists on a subset of the state space is empirically unproblematic. This is because this subset is independent of the parameters of the empirical model and is defined by a combination of choices and states that are observable. Hence the estimation can be conditioned on the subset of the data with observed $k' > k$ and $r > 0$, as the function $\widetilde{\omega}(k', k)$ exists on that subset and the unobserved productivity state can be recovered independently of the parameters of the model. This key insight allows us to turn to the empirical work.

2.3 Estimation

The strict monotonicity of the optimal capital choice in current productivity in the presence of R&D in Theorem 2.4 allows us to follow along the lines of Olley & Pakes (1996) to obtain estimates of production function parameters and of the unobserved productivity states.

The key insight in this section is that, if the true underlying model includes R&D, one has to control for the effect of R&D on the future productivity distribution when estimating the coefficients of the quasi-fixed factors. This raises the question of whether the estimation approach of Olley & Pakes (1996) needs to be modified. We present an alternative to the estimation specification in Olley & Pakes (1996) that addresses the effect of R&D directly. Due to monotonicity properties of the model, we can show that the original Olley-Pakes approach is asymptotically equivalent to this alternative. This implies that, in the context of the theoretical model in this study, the original Olley-Pakes method is valid even in the presence of R&D – as is our alternative.

However, while the asymptotic equivalence result in this section is specific to the dynamic model, the problem is more general: The problem is present whenever the expected future productivity depends on R&D (i.e. R&D has some effect) and is independent of whether one uses investment to proxy for productivity differences (Olley & Pakes 1996) or whether alternative proxies such as intermediate inputs are employed as proposed by Levinsohn & Petrin (2003). The intermediate input approach in Levinsohn & Petrin (2003) is attractive because it does not require a monotonicity proof for a structural dynamic model to derive a proxy for unobserved productivity. However, if the underlying model includes R&D, one still has to write down the candidate structural model and explore how to adequately correct the estimation procedure to yield consistent estimates for the quasi-fixed inputs.

This main point will now be developed in more detail. First, the production function is introduced and expected biases that arise if one fails to control for unobserved productivity differences across firms are discussed. We then present different variants of the estimation algorithm originally proposed by Olley & Pakes (1996), that apply in different settings with or without a selection problem and with or without R&D.

2.3.1 Production function and OLS biases

Following Olley & Pakes (1996), assume that firms in the industry produce a homogeneous product using a Cobb-Douglas production technology and that productivity differences result in Hicks neutral efficiency differences across firms i and time t,

$$y_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it} + \eta_{it} \tag{2.5}$$

where y represents the log of a measure of output (value added or sales), l the log of the labour input (e.g. number of employees), k the log the capital stock, ω the productivity state (unobserved by the econometrician but observed by the firm), and η an error term which represents either a serially uncorrelated additional productivity shock or measurement error (which can be serially correlated). Labour l is assumed to be a completely variable factor of production (i.e. not a state variable in the underlying model), while the capital stock in period t was chosen in the previous period through the investment decisions according to the capital accumulation equation[3]. To simplify notation, we will drop the subscript i in the following discussion but output, input factors, productivity and the error terms are understood to be firm and time specific whereas the production function coefficients are constant across time and firms.

The timing of the investment and input decisions is important. In line with the theoretical model, we assume that at the beginning of period t, firm i observes its productivity state ω_t and capital stock k_t. If the firm decides to continue operations, it decides how much of the variable factor labour to employ and chooses the levels of investment in physical capital and in R&D. The additional shock η_t is realised only after these choices are made. So while the input choice l_t responds to and is hence correlated with the productivity state ω_t, l_t is uncorrelated with the error term η_t. The same applies for the capital stock k_t. Even though capital k_t is predetermined by last periods capital stock and investment, it is correlated with ω_t to the extent that ω_t is correlated with ω_{t-1}, as the investment decision in $t-1$ was made on the basis of the distribution of ω_t conditional on ω_{t-1} (and R&D expenditure in $t-1$). The assumptions on η ensure that k_t is uncorrelated with η_t or previous η's.

The fact that unobserved productivity ω_t is correlated with the inputs l_t and k_t results in two well known biases in the OLS estimates of the parameters in equation (2.5) when one does not control for unobserved productivity ω. The first is due to the endogeneity of the input choices and has been recognised at least as early

[3]With slight abuse of notation, lower case letters now denote the logs of variables. The only instance where this makes a difference is for the capital accumulation equation. With lower case letters denoting the logs of the respective variables, the capital accumulation equation should read $\exp(k') = (1-\delta)\exp(k) + \exp(i)$.

capital fixed or sunk cost ?

as Marschak & Andrews (1944). Conditional on capital, more productive firms have a higher marginal product of labour in (2.5) and will, therefore, employ more labour. Similarly, highly productive firms will invest more in physical capital and so the next period's capital stock is positively correlated with productivity ω. Levinsohn & Petrin (2003) show that in the two input case, this endogeneity leads to an upward bias in the OLS estimate of the labour coefficient and to a downward bias in the estimate of the capital coefficient, provided there is a higher contemporaneous correlation between labour and productivity than between capital and productivity and provided labour and capital are positively correlated in the data – conditions which are likely to be satisfied.

The second bias is due to the selection of firms through exit and has been discussed in the literature at least since Mansfield (1962) and Wedervang (1965). The form of the exit rule (2.2) implies that a firm optimally decides to exit if its productivity state falls below the exit threshold which is a function of the capital stock of the firm (i.e. firm size). If the exit threshold is decreasing in capital (which will be the case if profits are strictly increasing in capital), the lower bound on the range of productivity realisations for the surviving firms that is observed in the data is decreasing in capital. If this translates into an average productivity among the survivors that is decreasing in the capital stock (which is not necessarily the case in the model with R&D), this leads to a downward bias in the capital coefficient. explain

2.3.2 Estimation approach

Numerous approaches to overcome the biases in the production function estimates have been proposed in the literature most notably OLS, fixed effects, and the Blundell & Bond (2000) instrumental variables approach (see Olley & Pakes (1996) and Levinsohn & Petrin (2003) for a discussion). The virtue of the estimation method proposed by Olley & Pakes (1996) is that it overcomes the problem of firm specific, time varying unobserved productivity in the estimation of production functions. Their method crucially relies on the monotonicity of the investment policy function in unobserved productivity (conditional on current capital) of the underlying structural dynamic model. The key insight is that this policy function can be inverted to express unobserved productivity as a function of the capital stock and investment (or, equivalently, as a function of capital at t and capital at $t+1$). The Olley-Pakes estimation algorithm exploits this fact in two stages. The first stage yields a consistent estimate of the production function coefficients of the variable factors of production (labour). The second stage estimates the coefficients of the quasi-fixed inputs.

The estimation approach followed in this study is almost identical to the Olley-Pakes algorithm. In fact, the key theoretical result in Corollary 2.5 of the invertibility of the investment policy function in the model with R&D implies that the first stage of the Olley-Pakes algorithm can be directly adopted to yield coefficient estimates for the variable factor(s) of production. However, the fact that the distribution of future productivity depends not only on current productivity but also on the amount of R&D investment requires a careful analysis of the second stage of the Olley-Pakes algorithm which estimates the parameters of the quasi-fixed inputs. We develop an alternative approach for the stage two estimation to address this issue and show that it is asymptotically equivalent to the original specification in Olley & Pakes (1996).

Stage one: Estimation of the coefficients of the variable input(s)

The first stage of the estimation approach is identical to the Olley-Pakes method and yields estimates of the coefficients of the variable factor(s) of production (in our case labour). The estimation strategy here is to control for the unobserved productivity nonparametrically exploiting the monotonicity property of the investment policy function.

According to Corollary 2.5, unobserved productivity can be expressed as a function of the current and future capital stock

$$\omega_t = \widetilde{\omega}(k_{t+1}, k_t), \tag{2.6}$$

where the functional form of $\widetilde{\omega}(\cdot, \cdot)$ is unknown and depends in a complex way on all the primitives of the structural model. Substituting (2.6) in the production function (2.5) and rewriting yields:

$$\begin{aligned} y_t &= \alpha_l l_t + \phi(k_{t+1}, k_t) + \eta_t, \text{ where} \\ \phi_t &= \phi(k_{t+1}, k_t) \equiv \alpha_0 + \alpha_k k_t + \widetilde{\omega}(k_{t+1}, k_t). \end{aligned} \tag{2.7}$$

With the functional form of $\widetilde{\omega}(\cdot, \cdot)$ and hence of $\phi(\cdot, \cdot)$ unknown, equation (2.7) is a partially linear semiparametric model (Robinson 1988). Semiparametric estimation of this equation yields an estimate of α_l and estimates of the unknown function $\phi(\cdot, \cdot)$. Note that for the identification of α_l, some variation in l_t that is uncorrelated to k_{t+1} and k_t (and hence ω_t) is required. This variation could e.g. be due to (serially uncorrelated) shocks in the wage rates.[4]

Stage two: Estimation of the coefficients of the quasi-fixed input(s)

The second stage of the estimation recovers the capital coefficient. The difficulty lies in separating the contribution of capital k_t to the term ϕ_t in equation (2.7) from the contribution of ω_t. The idea to identify α_k is that the current period's capital stock was chosen in $t-1$ and will therefore only be correlated with the expectation of productivity based on information available in $t-1$ but not with the productivity innovation. Identifying α_k then hinges on whether one can control for expected productivity conditional on past information. The correct way to control for this expectation depends on whether self-selection of firms through exit is a concern and whether a model with or without R&D is considered. The general strategy for stage two is again identical to the approach in Olley & Pakes (1996), but we will depart from their approach in the way we control for expected productivity.

Rearranging the production function (2.5) to define a transformed dependent variable y^* yields

$$y_t^* \equiv y_t - \alpha_l l_t = \alpha_0 + \alpha_k k_t + \omega_t + \eta_t.$$

The expectation of y_t^* conditional on information at $t-1$ and survival until t is

[4]With production function $f(\omega_t, k_t, l_t)$, the optimal labour input conditional on capital, productivity and the real wage w_t is $l(\omega_t, k_t, w_t) = \arg\max_{l_t} f(\omega_t, k_t, l_t) - w_t \exp(l_t)$. So random shocks to the real wage will lead to fluctuations in the labour input that are uncorrelated with ω_t and k_t and uncorrelated with the error terms in equation (2.7)

then

$$E[y_t^*|I_{t-1}, \chi_t = 1] = \alpha_0 + \alpha_k k_t + E[\omega_t|\psi_t, \chi_t = 1], \tag{2.8}$$

where I_{t-1} denotes the information set in $t-1$ and where the choice of distribution ψ_t in $t-1$ is sufficient to characterise the distribution of ω_t by assumption. Since the Markov assumption for the productivity process implies that productivity conditional on survival can be rewritten as $\omega_t = E[\omega_t|\psi_t, \chi_t = 1] + \xi_t$, the second stage estimation equation becomes

$$y_t^* = \alpha_0 + \alpha_k k_t + E[\omega_t|\psi_t, \chi_t = 1] + \xi_t + \eta_t, \tag{2.9}$$

where the productivity innovation ξ_t is uncorrelated with k_t. To estimate α_k consistently from this equation, we have to control for the expected productivity conditional on survival. Since the expectation term is again an unknown function we will have to take a similar approach as in stage one and control for the expectation nonparametrically. The key feature of the model with R&D is that $F(\cdot|\cdot)$ depends on the choice variable $\psi_t = \widetilde{\psi}(\omega_{t-1}, k_{t-1})$ and therefore on ω_{t-1} and k_{t-1} whereas in the Olley-Pakes model without R&D, this distribution is a function of the state ω_{t-1} only (i.e. $\psi_t = \omega_{t-1}$ in our notation). This dependence of the expectation on k_{t-1} in the model with R&D introduces an additional problem in the identification of α_k.

No selection

To illustrate this and to keep this point transparent, assume for the moment that self-selection of firms through exit is not an issue.[5] The expectation term in (2.9) then becomes

$$\begin{aligned} E[\omega_t|\psi_t, \chi_t &= 1] = E[\omega_t|\psi_t] = \int \omega' dF(\omega'|\psi_t) \\ &\equiv g(\psi_t) - \alpha_0. \end{aligned} \tag{2.10}$$

[5]The results reported in Levinsohn & Petrin (2003), for example, abstract form selection issues although the authors report that their results are insensitive to the inclusion of a selection stage.

For reasons that become apparent below, the function $g(\psi_t)$ in this equation is defined to capture the expectation term and the constant in equation (2.9).

No R&D: In the Olley-Pakes model without R&D, the firm cannot influence the distribution of productivity, i.e. $\psi_t = \omega_{t-1}$. As a result, the second stage estimation equation without R&D and without selection becomes

$$y_t^* = \alpha_k k_t + g(\omega_{t-1}) + \xi_t + \eta_t. \tag{2.11}$$

Since $\omega_{t-1} = \phi_{t-1} - \alpha_k k_{t-1} - \alpha_0$ by equation (2.7), this equation can be estimated semiparametrically where the unknown function $g(\omega_{t-1})$ can be approximated by a nonparametric function in $\phi_{t-1} - \alpha_k k_{t-1}$. Note that in this estimation equation α_k enters both the linear term $\alpha_k k_t$ and the nonparametric function in $\phi_{t-1} - \alpha_k k_{t-1}$, making the right hand side of the equation nonlinear in α_k.

R&D: In the context of a model with R&D, ψ_t is a choice which depends on the states in $t-1$ via the policy function (2.4). Identifying α_k in this case is not as straightforward: In particular, α_k is not identified if we approximate the function $g(\widetilde{\psi}(\omega_{t-1}, k_{t-1}))$ in (2.10) by a nonparametric function in $\phi_{t-1} - \alpha_k k_{t-1}$ and k_{t-1}. This is because k_t itself is an unknown function of ω_{t-1} and k_{t-1} via the policy function (2.3), implying that the entire right hand side of (2.8) is an unknown function of ϕ_{t-1} and k_{t-1}. To solve this identification problem we need to find an alternative way to proxy for ψ_t without having to use k_{t-1}. Two approaches suggest themselves:

The first uses data on R&D investments. The R&D function $r(\psi_t, \omega_{t-1})$ of the model is increasing in ψ_t for fixed ω_{t-1} (because a better distribution requires higher R&D expenditure), so R&D can be inverted to yield

$$\psi_t = r^{-1}(r_{t-1}, \omega_{t-1}), \tag{2.12}$$

where r_{t-1} denotes the observed R&D spending of firm i in period $t-1$. Using equation (2.12) to control for the distribution in period t, the second stage estimation equation then becomes

$$
\begin{aligned}
y_t^* &= \alpha_k k_t + g(r^{-1}(r_{t-1}, \omega_{t-1})) + \xi_t + \eta_t \qquad (2.13) \\
&= \alpha_k k_t + \widetilde{g}(r_{t-1}, \phi_{t-1} - \alpha_k k_{t-1}) + \xi_t + \eta_t.
\end{aligned}
$$

Equation (2.13) can now be used to obtain estimates for α_k replacing $g(r^{-1}(\cdot,\cdot))$ by a nonparametric function $\widetilde{g}(\cdot,\cdot)$ in $\phi_{t-1} - \alpha_k k_{t-1}$ and r_{t-1}. This approach is, of course only available if R&D data is available. However, it also requires that lagged R&D be uncorrelated with the error terms in (2.13). This can be violated if R&D is used in the construction of the value added measure y as is the case with our data (see below).

The second approach does not rely on R&D data but exploits another property of the structural model in the previous section, namely that the choice of distribution ψ_t in $t-1$ can be expressed as a function of the optimal choice k_t and the state variable ω_{t-1}. This is because conditional on the optimal capital choice in $t-1$, k_t, the choice of distribution in $t-1$, ψ_t, only depends on the productivity state ω_{t-1}:

$$
\psi_t = \overline{\psi}(\omega_{t-1}, k_t) = \arg\max_{\psi} -r(\omega_{t-1}, \psi) + \beta EV(\omega_t, k_t|\psi). \qquad (2.14)
$$

The dependence of ψ_t on k_{t-1} in the policy function therefore arises only indirectly through the link of k_t to k_{t-1}. Using this fact, the stage two estimation equation of the R&D model without selection becomes

$$
\begin{aligned}
y_t^* &= \alpha_k k_t + g(\overline{\psi}(\omega_{t-1}, k_t)) + \xi_t + \eta_t \\
&= f(\phi_{t-1} - \alpha_k k_{t-1}, k_t) + \xi_t + \eta_t. \qquad (2.15)
\end{aligned}
$$

This equation is no longer a partially linear semiparametric equation, but is now "fully" nonlinear. However, we can identify α_k from this equation even though $f(\cdot,\cdot)$ is an unknown function. Identification of α_k now purely comes the fact that the parametric specification of the production function provides the functional form of the first argument for the unknown function $f(\cdot,\cdot)$.

Three remarks are in order:

(1) Equation (2.15) is a multiple index model with only one index parameter, $-\alpha_k$, to be estimated and the others restricted to zero or unity in the obvious manner.

(2) Note that this approach does not require the production function to be Cobb-Douglas. All that is required is that ω_{t-1} is a parametric function of ϕ_{t-1} and the quasi fixed factors and that the parameter(s) of interest can be identified from this parametric function.

(3) Similar to the monotonicity proof which is the basis for the productivity proxy, this approach to identification of the parameters of the quasi-fixed factors depends delicately on the structure of the model in the previous section. It relies on the equation (2.14), i.e. on the fact that the choice of distribution can be expressed as a function of productivity state and the capital *choice* rather than the capital *state*. This property of the model is not necessarily robust to even slight modifications of the setup. So even if alternative ways to proxy for productivity are available that do not require a full structural model (e.g. intermediate inputs as in Levinsohn & Petrin (2003)), one still needs to write down a structural model that makes explicit how R&D affects the expected productivity and therefore justifies the approach to proxy for the expectation (i.e. equation (2.14), equation (2.12), or a similar equation).

So when selection is not an issue and the true model is our model with R&D, there are two possibilities: When R&D data is available and uncorrelated with the error terms, the system (2.7) and (2.13) can be estimated using the first stage estimates of α_l and ϕ_t in the second stage. Alternatively the system (2.7) and (2.15) can be estimated again using first stage estimates in stage two. The second alternative does not require R&D data. In both cases, estimation in two stages is necessary as the innovation in productivity, ξ_t, is correlated with l_t so that α_l could not be estimated consistently from equation (2.13) or (2.15).

Selection

Now consider the case where self-selection of firms through exit is of concern. In this case, the expectation of productivity conditional on past information and survival becomes

$$
\begin{aligned}
E[\omega_t|\psi_t,\chi_t &= 1] = \frac{\int\limits_{\omega'\geq\underline{\omega}_t} \omega' dF(\omega'|\psi_t)}{\int\limits_{\omega'\geq\underline{\omega}_t} dF(\omega'|\psi_t)} \\
&= [\Pr(\chi_t = 1|\underline{\omega}_t,\psi_t)]^{-1} \int\limits_{\omega'\geq\underline{\omega}_t} \omega' dF(\omega'|\psi_t) \\
&\equiv g(\psi_t,\underline{\omega}_t) - \alpha_0. \qquad (2.16)
\end{aligned}
$$

Again, the function $g(\psi_t,\underline{\omega}_t)$ is defined to capture the expectation term and the constant in the production function. The only difference between equation (2.16) and the expression in Olley&Pakes (1996) is the choice ψ_t which replaces the past state ω_{t-1} in their paper.

No R&D: Let us first discuss the Olley-Pakes approach to control for selection in the model without R&D. To obtain a proxy for the second index $\underline{\omega}_t$ Olley & Pakes exploit a separate estimate of the survival probability. Rewriting the survival probability yields

$$
\begin{aligned}
\Pr(\chi_t &= 1|\underline{\omega}_t,\psi_t) = \Pr(\widetilde{\chi}(\omega_t,k_t) = 1|\underline{\omega}(k_t),\widetilde{\psi}(\omega_{t-1},k_{t-1})) \\
&= \widetilde{p}(k_t,k_{t-1}) \equiv P_t,
\end{aligned}
$$

where the second line follows because ω_{t-1} is a function of k_t and k_{t-1} by equation (2.6). Note also that k_t is chosen in $t-1$ and is observable in $t-1$ (through investment i_{t-1} and k_{t-1}). Estimates for the survival probabilities P_t, can therefore be obtained by regressing survival in t on k_t and k_{t-1} in a suitably flexible way.

Provided the density of ω_t (conditional on ψ_t) has positive support around $\underline{\omega}_t$ the probability of survival will be strictly decreasing in the exit threshold $\underline{\omega}_t$. This in turn implies that the survival probability P_t can be inverted to obtain

$$\underline{\omega}_t = f(\psi_t, P_t)$$

which forms the basis to control for $\underline{\omega}_t$ in (2.16). Combining these facts, the second stage estimation equation as proposed by Olley & Pakes becomes

$$
\begin{aligned}
y_t^* &= \alpha_k k_t + g(\psi_t, f(\psi_t, P_t)) + \xi_t + \eta_t \\
&= \alpha_k k_t + \widetilde{g}(\phi_{t-1} - \alpha_k k_{t-1}, P_t) + \xi_t + \eta_t. \qquad (2.17)
\end{aligned}
$$

because $\psi_t = \omega_{t-1}$ in their model and because $\omega_{t-1} = \phi_{t-1} - \alpha_k k_{t-1} - \alpha_0$. This equation is again a partially linear semiparametric model from which α_k can be estimated.

R&D: Controlling for ψ_t in the model with R&D is slightly more complicated. As before we can express ψ_t either as a function of ω_{t-1} and r_{t-1} through (2.12) or as a function of ω_{t-1} and k_t through (2.14). In what follows, we will only discuss the second possibility. [6]

Note that the form of the optimal policy and the assumptions on the policy function imply that the exit threshold is a function of k_t, i.e. $\underline{\omega}_t = \underline{\omega}(k_t)$. Combining this with equation (2.14), we can rewrite the second stage estimation equation as

$$
\begin{aligned}
y_t^* &= \alpha_k k_t + g(\overline{\psi}(\omega_{t-1}, k_t), \underline{\omega}(k_t)) + \xi_t + \eta_t \\
&= f(\phi_{t-1} - \alpha_k k_{t-1}, k_t) + \xi_t + \eta_t. \qquad (2.18)
\end{aligned}
$$

This estimation equation for the R&D model with selection is identical to the estimation equation (2.15) for the model without R&D. This implies that by using equation (2.14) to control for ψ_t one also implicitly controls for selection. Both cases lead to the same "fully" nonlinear estimation equation (2.18).[7]

To understand why this is the case, consider the following alternative interpretation of equation (2.18). Since productivity and capital are the state variables of the model, output in the current period can be expressed as expected output

[6] Deriving an estimation equation using the first option is straightforward. Following Olley and Pakes' approach to control for $\underline{\omega}_t$, we could e.g. rewrite $g(\psi_t, \underline{\omega}_t) = \widetilde{g}(r_{t-1}, P_t, \phi_{t-1} - \alpha_k k_{t-1})$, which can used to estimate the parameters of interest.

[7] Of course, the functions $f()$ and $g()$ are different in each case but always unknown. So the estimation equations are identical in terms of the arguments entering the unknown functions and will therefore yield identical estimation results for a given estimation technique.

conditional on the lagged state variables plus an error term which is uncorrelated with the lagged state. To identify α_k, we have shown that instead of conditioning on the lagged capital state, we can also condition on the lagged capital choice, i.e. on current capital. This gives rise to the multiple index model in equation (2.18).[8]

So whether selection is a problem or not, production function coefficients in the model with R&D can be estimated from the system in (2.7), and (2.18). Stage one yields estimates for α_l and ϕ_t from (2.7) which can then be substituted for the true values in the estimation of (2.18). Again, estimation in two stages is necessary, because ξ_t is correlated with l_t.[9]

Asymptotic equivalence of the Olley-Pakes approach and the fully nonlinear specification:

While the motivations for the estimation equation (2.17) in Olley & Pakes and equation (2.18) for the R&D model with selection are different, they are similar in the sense that in each equation the nonparametric function has two arguments. The question is whether the inclusion of an estimate of the survival probability in equation (2.17) in the Olley & Pakes version is also sufficient to control for the joint effect of R&D and selection. In this case, estimates from the two approaches should be asymptotically identical.

This will clearly be the case if the survival probability can be inverted to yield k_t as a function of ω_{t-1} and P_t. Recall that the survival probability is strictly increasing in the distribution ψ_t and strictly decreasing in the exit threshold $\underline{\omega}_t$. The assumption on the profit function imply that $\underline{\omega}(k_t)$ is nonincreasing in k_t. Furthermore, an Euler equation argument for the optimal choice of ψ_t conditional on ω_{t-1} and k_t shows that $\overline{\psi}(\omega_{t-1}, k_t)$ is strictly increasing in k_t. Together, these two monotonicity properties imply that the inversion of the survival probability is

[8]It is interesting to note that this rationale not only holds for output, but for all variables (labour, investment etc.) in the model. While this study does not pursue this any further, additional estimation equations for these variables could be added to stage two of the estimation to improve efficiency. Since these additional equations would all be nonparametric multiple index models of the same indices, no additional assumptions on the model would be required.

[9]The same remarks as in the no selection case apply.

possible. Therefore, the two approaches are asymptotically equivalent and using a nonparametric function in ω_{t-1} and P_t to proxy for the expected productivity conditional on survival should asymptotically yield the same results as a nonparametric function in ω_{t-1} and k_t even in our model with R&D.[10]

In practice, whether the two approaches produce similar estimates for the coefficients of the quasi-fixed factors is an empirical question.

2.4 Data

We use firm level data over the period 1980-2001 for four "3-digit" SIC industries. The industries under study are "Pharmaceuticals (SIC 283)", "Computer Hardware (SIC 357)", "Telecommunications Equipment (SIC 366)", and "Software (SIC 737)". The data is an unbalanced panel constructed from the COMPUSTAT database. COMPUSTAT contains accounting and financial market data and covers publicly traded companies on North American stock markets that submit reports to the Securities and Exchange Commission (SEC).

Table 2.1 reports the number of firms and number of firm-year observations by 3-digit industry and constituent 4-digit industry included in the estimation sample. With almost 700 firms and 4550 observations, "Software" is the biggest industry in our sample followed by "Pharmaceuticals" with 461 firms, "Telecom Equipment" with 253 firms and "Computers" with 259 firms.[11] Table 2.2 lists the 10 biggest firms in each industry in terms of sales revenue in 2001. The prominent suspects appear on this list.

For the estimation sample, we only include firms that contribute at least one observation to stage 2 of the estimation, i.e. firms for which at least one lag is available. We also exclude all firm year observations with negative or zero reported investment as the monotonicity proof of the investment policy function is not available for these observations.[12] Since our estimation equations are in logs,

sell assets

[10] An advantage of the fully nonparametric approach in equation (2.18) is that its logic directly carries over to a model with two or more quasi-fixed inputs (such as separate capital stocks for property and equipment).

[11] Details on how these industries and the constituing 4-digit industries were selected are given in Appendix A2.2.

[12] As in Olley & Pakes (1996), weak monotonicity of the policy function holds everywhere, while

Table 2.1: Number of firms and firm years by 3-digit and 4-digit SIC industry

a. Pharmaceuticals (SIC 283)

	SIC	# firms	# firm years
PHARMACEUTICAL PREPARATIONS	2834	245	2245
IN VITRO,IN VIVO DIAGNOSTICS	2835	87	781
BIOLOGICAL PDS,EX DIAGNSTICS	2836	129	975
Total		**461**	**4001**

b. Computer Hardware (SIC 357)

	SIC	# firms	# firm years
ELECTRONIC COMPUTERS	3571	73	617
COMPUTER STORAGE DEVICES	3572	58	534
COMPUTER TERMINALS	3575	28	242
COMPUTER PERIPHERAL EQ, NEC	3577	94	849
Total		**253**	**2242**

c. Telecom Equipment (SIC 366)

	SIC	# firms	# firm years
TELE & TELEGRAPH APPARATUS	3661	134	1129
RADIO, TV BROADCAST, COMM EQ	3663	125	1321
Total		**259**	**2450**

d. Software (SIC 737)

	SIC	# firms	# firm years
COMPUTER PROGRAMMING SERVICE	7371	38	302
PREPACKAGED SOFTWARE	7372	606	3842
CMP PROCESSING,DATA PREP SVC	7374	49	406
Total		**693**	**4550**

Table 2.2: Biggest firms in terms of sales revenue 2001 by industry

Company	SIC	Sales 2000 (billion US$)
a. Pharmaceuticals (SIC 283)		
MERCK & CO	2834	40.36
PFIZER INC	2834	29.57
JOHNSON & JOHNSON	2834	29.14
GLAXOSMITHKLINE PLC	2834	27.27
NOVARTIS AG	2834	22.10
AVENTIS SA	2834	21.43
ROCHE HOLDINGS LTD	2834	18.78
BRISTOL MYERS SQUIBB	2834	18.22
PHARMACIA CORP	2834	18.14
ASTRAZENECA PLC	2834	15.80
b. Computer Hardware (SIC 357)		
NEC CORP	3571	42.93
COMPAQ COMPUTER CORP	3571	42.38
DELL COMPUTER CORP	3571	31.89
CANON INC	3577	24.19
XEROX CORP	3577	18.70
SUN MICROSYSTEMS INC	3571	15.72
GATEWAY INC	3571	9.60
EMC CORP/MA	3572	8.87
APPLE COMPUTER INC	3571	7.98
SEAGATE TECHNOLOGY	3572	6.45
c. Telecom Equipment (SIC 366)		
MOTOROLA INC	3663	37.58
NORTEL NETWORKS CORP	3661	30.29
ALCATEL	3661	29.49
ERICSSON (L M) TEL	3663	29.22
NOKIA CORP	3663	28.52
SHARP CORP	3663	16.36
MARCONI PLC	3661	9.85
THOMSON	3663	8.54
AVAYA INC	3663	7.68
TELLABS INC	3661	3.39
d. Software (SIC 737)		
MICROSOFT CORP	7372	22.96
ORACLE CORP	7372	10.86
AUTOMATIC DATA PROCESSING	7374	8.37
FIRST DATA CORP	7374	5.71
COMPUTER ASSOCIATES INTL INC	7372	4.20
COMPUWARE CORP	7372	2.01
SIEBEL SYSTEMS INC	7372	1.80
PEOPLESOFT INC	7372	1.74
SUNGARD DATA SYSTEMS INC	7372	1.66
FISERV INC	7374	1.65

we will also loose all observations with negative value added estimates. Finally, some companies grow very rapidly through acquiring other firms. Even though this may be a natural way to grow, the small number of firm years for which the sales contribution of acquisitions in the exceed 50% of sales is excluded. Table 2.3 reports the impact of this data cleaning exercise in terms of loss of observations.

Table 2.4 shows summary statistics for the variables in the resulting estimation sample by industry.[13] In all industries the mean and median levels of R&D investments exceeds the corresponding levels in physical capital investments. This suggests that R&D is, in fact, an important choice variable in these industries and that abstracting from R&D in the underlying dynamic model of the industry may be problematic. While "Pharmaceuticals" has the highest average and median capital stock, and the highest levels of R&D investments, "Computers" and "Telecom Equipment" have the highest levels of employment.

Firms exiting the dataset are assigned a reason for deletion by COMPUSTAT. The main exit reasons by number of occurrences are "merger and acquisition", "bankruptcy", "liquidation", and a category named "other reasons" which includes companies that have stopped reporting to the SEC.[14] This last category is most prominent for small companies. For our purposes, exit reasons are grouped into three categories: "merger and acquisition", "bankruptcy and liquidation", and "other". Table 2.5 shows the number of survivors and exiting firms in our sample for each 3-digit industry and by exit reason. 54% of the firms in our sample survive, 33% exit through mergers, 3.5% go bankrupt or liquidate and 9% exit for other reasons. While bankruptcy and liquidation are clearly events indicating the failure of a company, "merger and acquisition" can be a success or a failure. We believe that the reason "other" is also a failure category.

---

strict monotonicity which is required for the inverstion can be guaranteed only for the subset of the data with positive investments.

[13]Details on the construction of variables and the price deflators used are given in Appendix A2.2. Note that industry specific output and investments deflators are used for "Pharmaceuticals", "Computers", and "Telecom Equipment". Since we lack consistent deflators for "Software" over this period, the GDP deflator is used for this industry.

[14]Firms with fewer than 10 million US$ in assets or fewer than 500 shareholders submit reports to the SEC voluntarily.

Table 2.3: Number of observations dropped by industry

	SIC	# raw obs	VA<=0	Invmt<=0 or missing	Sales contrib. of acq.>.5
Pharmaceut.	**283**	5078	862	132	193
Computers	**357**	2617	307	21	56
Telecom Eqmt	**366**	2814	307	24	25
Software	**737**	5479	696	113	96

Table 2.4: Summary statistics by industry for estimation sample

a. Pharmaceuticals (SIC 283)

	# obs	mean	std. dev	median	min	max
Employment	4001	4.69	13.95	0.16	0.0010	116.18
Capital	4001	330.04	1110.27	11.43	0.0015	13731.12
Value Added	4001	541.82	1953.52	11.00	0.0035	23428.58
Investment	4001	68.35	226.04	1.93	0.0009	2477.87
R&D	3705	117.75	353.64	9.86	0.0057	4896.81

b. Computer Hardware (SIC 357)

	# obs	mean	std. dev	median	min	max
Employment	2242	4.91	17.67	0.39	0.0020	157.77
Capital	2242	250.09	1210.47	8.57	0.0300	18193.68
Value Added	2242	272.88	973.31	18.37	0.0122	9895.70
Investment	2242	64.08	283.52	2.23	0.0008	3646.04
R&D	2175	92.54	589.96	2.79	0.0002	11350.19

c. Telecom Equipment (SIC 366)

	# obs	mean	std. dev	median	min	max
Employment	2450	4.78	20.17	0.35	0.0020	213.10
Capital	2450	206.45	1066.28	9.68	0.0337	17915.55
Value Added	2450	353.44	1668.93	15.15	0.0035	17933.26
Investment	2450	54.96	302.47	1.87	0.0010	7217.81
R&D	2311	84.71	448.62	3.29	0.0011	7166.13

d. Software (SIC 737)

	# obs	mean	std. dev	median	min	max
Employment	4550	1.14	3.53	0.27	0.0030	47.60
Capital	4550	48.08	172.61	8.76	0.0043	2909.02
Value Added	4550	122.68	657.54	18.79	0.0052	21257.28
Investment	4550	11.45	44.97	1.71	0.0010	1008.07
R&D	3858	27.83	131.34	6.53	0.0016	4002.13

Units:
Employment: 1000 employees
Other variables: million (1996) US$

Table 2.5: Survivors and exiting firms by industry in estimation sample

	SIC	Survivors	Exit by reason			Total	Total
			M&A	Bkr/Lqn	Other	Exits	
Pharmaceut.	**283**	291	137	7	26	170	461
Computers	**357**	99	89	30	35	154	253
Telecom Eqmt	**366**	127	87	11	34	132	259
Software	**737**	375	247	14	57	318	693
Total		892	560	62	152	774	1,666

2.5 Production function estimates

2.5.1 Specification details and standard errors

The estimation strategy for the system in (2.7) and (2.18) (respectively its variants) relies on obtaining nonparametric estimates of the unknown functions $\phi()$ and $f()$ (respectively $\phi()$, $p()$, and $g()$) and use them as if they were their true counterparts in the estimation of (2.18). The first order condition with respect to α_k for a nonlinear least squares estimator of (2.18) falls in the class of moment conditions for which Pakes & Olley (1995) prove that a sufficiently smooth class of semiparametric estimators produce consistent, $\sqrt{n}$-consistent and asymptotically normal estimates. Olley & Pakes (1996) present estimation results using kernel estimates for the nonparametric functions as well as polynomial series approximations. While the latter are much easier to implement and are far less computationally intensive, only the kernel estimates fall within the class of sufficiently smooth semiparametric estimators for which the consistency and asymptotic normality proofs apply. However, the estimates of the two approaches in Pakes & Olley (1995) are remarkably close and the authors point out that while only their bias reducing kernel estimators "are known to abide by all the regularity conditions needed for [their] limit theorems [...] there is a strong presumption that the series estimates do also" (Pakes & Olley 1995, p. 329). Furthermore, in their study on Chilean industries, Levinsohn & Petrin (2003, footnote 27) also report that the series estimators produce very similar results to those produced by their kernel estimator.

On the basis of these experiences and because of the ease of implementation, we choose the polynomial series approach to approximate the unknown functions $\phi()$ and $f()$ (respectively $\phi()$ $p()$ and $g()$ depending on the estimation equation). In the stage one estimation of equation (2.7), we run an OLS regression of our output measure y_t on labour l_t and a polynomial series expansion in the variables proxying the joint effect of capital and productivity. As proxy variables we choose log investment i_t, capital k_t, and time t. Including i_t is clearly equivalent to including k_{t+1} as a proxy because of the deterministic capital accumulation equation. We choose i_t because it is less highly correlated with k_t than k_{t+1}. We also include a linear time variable t in the polynomial expansion to allow for changes in the policy function over time reflecting changes in the economic environment. Stage one results in estimates for $\widehat{\alpha}_l$ and $\widehat{\phi}_t$.

In stage two, we use a combination of grid search over $\widehat{\alpha}_k$ and derivative based optimization to minimise the sum of squared residuals in the relevant stage two regression equation. We approximate the unknown function $f()$ (respectively $g()$) by a polynomial series expansion in the relevant arguments. Where required, we run a probit regression of survival on a polynomial in the proxy variables to obtain an estimate of the survival probability $\widehat{P}_t$.

In all cases, we use a fifth order polynomial series expansion in stage one and a third order expansion in the probit regression and in stage two. We present estimates using third order expansion for stage two, because for higher order expansions discontinuities in the objective function in the parameter α_k arose in some cases.[15] As in Olley & Pakes (1996), neither the estimates nor the sums of squared residuals change much if the order of the polynomial expansions is increased.

Pakes & Olley (1995) also provide a method to compute asymptotic standard errors for production function estimates for the Olley-Pakes approach. Instead of adapting their formula to our procedure, we use a bootstrap approach with 200 repetitions to obtain standard errors. Besides being easy to implement, the bootstrap approach also provides us with new productivity estimates for each bootstrap draw. These will allow us to compute bootstrap standard errors in our analysis of the effect of R&D on the distribution of productivity in the next section that take into account that our productivity variables are estimates that change with each bootstrap sample.

Following Levinsohn & Petrin (2003), we resample from the population of firms with replacement, giving each firm equal probability to be selected. We continue resampling until the bootstrap sample contains the same or just exceeds the number of firm-year observations in stage two of the estimation as the original sample. Once a bootstrap sample has been constructed, it is held constant across specifi-

[15]Conditional on α_k the problem reduces to an OLS regression of the relevant dependent variable (which needs to be transformed if a linear term is present) on the series terms. The discontinuity of the objective function when expansions of order four or higher are used comes from the fact that the matrix $X'X$ is often close to singular in this case so that variables have to be dropped to compute the OLS estimates. For orders lower than four, we did not encounter this problem.

look at subsample of large firms – essentially no selection effect.

cations. As in Levinsohn & Petrin (2003), this makes testing whether estimates differ across specifications straightforward.

2.5.2 Results

long funnel ...

Production function estimates for the different specifications are presented in Table 2.6. Each panel corresponds to one industry, while the columns within each panel show estimates for different specifications. The first column (labelled "o") in each panel reports OLS estimates. The remaining columns correspond to the different variants of the Olley-Pakes estimation procedure discussed above. These specifications are labelled according to the variables that enter the nonparametric function of expected productivity in stage two of the estimation (in addition to lagged productivity which is always an argument of this function). Column "n" is the pure no selection, no R&D case where the expectation is only a function of lagged productivity (equation (2.11)). Column "r" is the no selection case with R&D corresponding to equation (2.13) where we proxy for the expectation using lagged R&D data. Column "k" is the fully non-linear specification in equation (2.18) that captures the effect of R&D and simultaneously also controls for selection in our model by putting the current capital stock in the nonparametric function. Finally, column "p" corresponds to the original Olley-Pakes stage 2 equation (2.17) that uses the lagged probability of survival. Since stage one is identical across the modified Olley-Pakes specifications, the labour coefficient is identical in columns "n" to "p" and we only report it for column "n". The standard deviations of the coefficient estimates across the 200 bootstrap samples are reported below the estimates.

As expected, the OLS point estimates of the labour coefficient are significantly higher than the labour coefficient in the Olley-Pakes approaches in all industries. In fact, this observation holds for each of the bootstrap samples across all industries except one repetition for "Pharmaceuticals".[16] Furthermore, the labour coefficient is estimated fairly precisely.

The discussion in the estimation section showed that we might expect a down-

[16] This point is apparent from Table 2.7, panel b which is discussed below.

Table 2.6: Production function estimates

a. Pharmaceuticals (SIC 283)

Dep. Var: Value Added		OLS	Modified Olley-Pakes procedure			
			Stage 2 variables in addition to lagged omega:			
			None	**L.R&D**	**Capital**	**L.P**
		o	n	r	k	p
Labour		**0.748**	**0.687**			
	SE	0.041	0.044			
Capital		**0.388**	**0.384**	**0.140**	**0.304**	**0.378**
	SE	0.038	0.070	0.077	0.114	0.073
# obs		3932	3269	3023	3269	3269
# firms		461	461	436	461	461
RSS		456.7	317.3	291.9	316.0	316.8

b. Computer Hardware (SIC 357)

Dep. Var: Value Added		OLS	Modified Olley-Pakes procedure			
			Stage 2 variables in addition to lagged omega:			
			None	**L.R&D**	**Capital**	**L.P**
		o	n	r	k	p
Labour		**0.973**	**0.833**			
	SE	0.037	0.043			
Capital		**0.067**	**0.189**	**0.078**	**0.206**	**0.291**
	SE	0.035	0.072	0.141	0.076	0.097
# obs		2210	1866	1821	1866	1866
# firms		253	253	245	253	253
RSS		186.9	139.8	134.3	139.4	138.5

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

All specifications other than OLS are labelled according to the variables included in stage 2 of the estimation procedure.
In all cases a 5th order polynomial expansion in investment, capital, and time is used to proxy for productivity in stage 1 and a 3rd order polynomial expansion is used in stage 2. The number of observations and RSS reported refer to stage 2.

The OLS specification also includes a 5th order polynomial in time to control for time effects. Including time dummies yields very similar results.

c. Telecom Equipment (SIC 366)

Dep. Var: Value Added	**OLS**	**Modified Olley-Pakes procedure** Stage 2 variables in addition to lagged omega:			
		None	**L.R&D**	**Capital**	**L.P**
	o	**n**	**r**	**k**	**p**
Labour	**0.826**	**0.688**			
SE	0.038	0.038			
Capital	**0.279**	**0.329**	**0.151**	**0.327**	**0.410**
SE	0.034	0.058	0.105	0.078	0.062
# obs	2443	2094	1976	2094	2094
# firms	259	259	255	259	259
RSS	203.0	158.2	143.0	158.0	156.6

d. Software (SIC 737)

Dep. Var: Value Added	**OLS**	**Modified Oliey-Pakes procedure** Stage 2 variables in addition to lagged omega:			
		None	**L.R&D**	**Capital**	**L.P**
	o	**n**	**r**	**k**	**p**
Labour	**0.775**	**0.684**			
SE	0.031	0.037			
Capital	**0.380**	**0.428**	**0.239**	**0.463**	**0.432**
SE	0.026	0.035	0.035	0.057	0.036
# obs	4526	3647	3089	3647	3647
# firms	693	693	631	693	693
RSS	387.5	239.4	192.4	239.1	238.8

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

All specifications other than OLS are labelled according to the variables included in stage 2 of the estimation procedure.
In all cases a 5th order polynomial expansion in investment, capital, and time is used to proxy for productivity in stage 1 and a 3rd order polynomial expansion is used in stage 2. The number of observations and RSS reported refer to stage 2.

The OLS specification also includes a 5th order polynomial in time to control for time effects. Including time dummies yields very similar results.

ward bias in the OLS capital coefficient. The regression results confirm this for the industries "Computers", "Telecom Equipment" and "Software", but not for "Pharmaceuticals" where, if anything, the OLS capital coefficient tends to be higher than the alternative estimates. The no selection, no R&D specification "n" produces estimates that seem more reasonable than the OLS estimates and that are fairly precisely estimated.[17]

The estimated capital coefficient drops drastically to unreasonable levels when we introduce lagged R&D to control for expected productivity in specification "r" (equation (2.13)). As mentioned in the estimation section, we suspect that this is due to an endogeneity problem with respect to R&D as the R&D data is unfortunately used in the construction of the value added measure (see Appendix A2.2).

In our model, the fully non-linear specification "k" in equation (2.18) controls for both, selection and R&D. Compared to OLS, this specification produces higher estimated capital coefficients for all industries except "Pharmaceuticals". Compared to the no selection, no R&D specification "n", the estimates are fairly close again, except for "Pharmaceuticals". Standard errors for this specification are higher than in specification "n" or the original Olley-Pakes specification "p". This is probably due to the fully non-linear specification as opposed to the partially linear specification.

In contrast to the results obtained by Olley & Pakes (1996), estimates from the original Olley-Pakes specification "p" differ from the no selection, no R&D estimates "n" for "Computer Hardware" and for "Telecom Equipment". However, the most puzzling observation is that coefficient estimates obtained from the original Olley-Pakes specification "p" seem to be quite different from those obtained from our specification "k". Except for "Software" where both yield similar results, the Olley-Pakes estimates for our samples are significantly higher than our "k" estimates. So despite the asymptotic equivalence of the two specifications, they can yield quite different point estimates.

To get a sense of the significance of the differences of the capital coefficient

[17]The results presented in Pakes & Olley (1995) suggest that, if anything, the bootstrapped standard errors are significantly higher than the asymptotic standard errors they derive.

across the specifications, Table 2.7 panel a reports the percentage of cases in which a particular inequality holds across the original and the bootstrap samples. The rows correspond to comparisons between two specifications and the columns to industries.

On this basis (and excluding specification "r"), the OLS capital coefficient is indistinguishable from those of the other specifications for "Pharmaceuticals" but is significantly lower for "Computers", "Telecom Equipment" and "Software". The exception is specification "k" which also yields similar estimates to OLS for "Telecom Equipment". Compared to the no selection, no R&D estimates "n", specification "p" yields significantly higher estimates for "Computers", while the estimates from specification "k" do not differ significantly from those of specification "n" on this basis. Finally, comparing specifications "p" and "k" on this basis gives little evidence that there is a significant difference between the two methods despite the sometimes substantial differences in the point estimates in Table 2.6.

Panel b of Table 2.7 shows that the difference in the labour coefficients between OLS and the invertibility approach are highly significant as discussed.

2.5.3 Specification tests

From these results alone, it cannot be concluded which of the specifications above are appropriate for each industry, if any. Olley & Pakes (1996) also propose a simple specification test. If the assumptions and specifications leading to the estimation equations are correct, the error term in the stage two estimation equation (2.9) should be mean independent from lagged observations of capital and labour. This is because the innovation ξ in productivity is uncorrelated to lags of these variables (as well as to current capital). Table 2.8 presents results for these specification tests. For each industry and for each of the specifications "n" (no selection, no R&D), "p" (the original Olley-Pakes specification), and "k" (the fully non-linear specification controlling for selection and R&D), the tables report regression results including lagged labour as a linear term in stage two. For specifications "n" and "p", it also reports regressions including a linear term in lagged capital in stage two.

For "Pharmaceuticals" the coefficients for these additional variables are very

Table 2.7: Differences in coefficient estimates across specifications

a. Capital coefficient

SIC	283	357	366	737
o > n	0.851	0.025	0.124	0.040
o > p	0.816	0.000	0.075	0.045
o > k	0.896	0.025	0.313	0.090
o > r	0.995	0.453	0.960	1.000
n > p	0.557	0.020	0.164	0.318
n > k	0.706	0.159	0.647	0.294
n > r	0.980	0.861	1.000	1.000
p > k	0.736	0.905	0.801	0.363
p > r	0.975	0.995	1.000	1.000
k > r	0.910	0.970	1.000	1.000

b. Labour coefficient

SIC	283	357	366	737
o >others	0.995	1.000	1.000	1.000

Each row of the table gives a comparison of the capital coefficient for two specifications. The number in each cell reports the percentage of cases across the 201 samples (one original sample + 200 bootstrapped samples) in which the capital coefficient in the first specification exceeds that of the second specification.
E.g. the first entry for row "p>k" implies that in 73.1% of the samples for Pharmaceuticals, the estimate of the capital coefficient in the Olley-Pakes specification "p" exceeds that in our fully nonlinear specification "k".

Table 2.8: Production function specification tests

a. Pharmaceuticals (SIC 283)

Dep. Var: Value Added

	n	n	p	p	k
Labour	**0.687**	**0.687**	**0.687**	**0.687**	**0.687**
SE	0.044	0.044	0.044	0.044	0.044
Capital	**0.386**	**0.379**	**0.381**	**0.381**	**0.301**
SE	0.068	0.064	0.071	0.073	0.112
L.Labour	**-0.002**		**-0.006**		**-0.020**
SE	0.011		0.011		0.016
fraction>0	0.746		0.612		0.174
L.Capital		**0.004**		**-0.005**	
SE		0.013		0.014	
fraction>0		0.896		0.751	

Numbers below coefficients are standard errors from 200 bootstrap samples and the percentage of cases with positive coefficients

b. Computers (SIC 357)

Dep. Var: Value Added

	n	n	p	p	k
Labour	**0.833**	**0.833**	**0.833**	**0.833**	**0.833**
SE	0.043	0.043	0.043	0.043	0.043
Capital	**0.214**	**0.194**	**0.319**	**0.299**	**0.217**
SE	0.073	0.071	0.103	0.096	0.070
L.Labour	**-0.020**		**-0.023**		**-0.107**
SE	0.007		0.011		0.035
fraction>0	0.000		0.005		0.000
L.Capital		**-0.009**		**-0.012**	
SE		0.004		0.007	
fraction>0		0.015		0.045	

Numbers below coefficients are standard errors from 200 bootstrap samples and the percentage of cases with positive coefficients

c. Telecom Eqmt (SIC 366)

Dep. Var: Value Added

	n	n	p	p	k
Labour	**0.688**	**0.688**	**0.688**	**0.688**	**0.688**
SE	0.038	0.038	0.038	0.038	0.038
Capital	**0.335**	**0.317**	**0.486**	**0.472**	**0.317**
SE	0.109	0.064	0.102	0.095	0.078
L.Labour	**-0.004**		**-0.035**		**-0.084**
SE	0.022		0.025		0.017
fraction>0	0.124		0.080		0.000
L.Capital		**0.008**		**-0.023**	
SE		0.008		0.021	
fraction>0		0.801		0.358	

Numbers below coefficients are standard errors from 200 bootstrap samples
and the percentage of cases with positive coefficients

d. Software (SIC 737)

Dep. Var: Value Added

	n	n	p	p	k
Labour	**0.684**	**0.684**	**0.684**	**0.684**	**0.684**
SE	0.037	0.037	0.037	0.037	0.037
Capital	**0.578**	**0.428**	**0.610**	**0.436**	**0.442**
SE	0.073	0.038	0.087	0.044	0.052
L.Labour	**-0.070**		**-0.087**		**-0.124**
SE	0.028		0.034		0.019
fraction>0	0.000		0.000		0.000
L.Capital		**0.000**		**-0.002**	
SE		0.007		0.009	
fraction>0		0.537		0.398	

Numbers below coefficients are standard errors from 200 bootstrap samples
and the percentage of cases with positive coefficients

close to zero, implying that all the specifications are accepted by this test. This is also confirmed by the fact that the coefficient estimates for labour and capital do not change very much once the lags are included. For "Computers" and "Software", the opposite emerges: Regardless of the specification, including lagged labour or capital always leads to significant negative coefficients for these variables, so that the tests reject these specifications. Finally, for "Telecom Equipment", we can accept the specifications "n" and "p", but the coefficient of lagged labour in specification "k" comes out significantly negative. This implies that the original Olley-Pakes specification cannot be rejected, while the fully non-linear specification is rejected by the data.

These tests suggests investment is a valid proxy for unobserved productivity differences across firms in the "Pharmaceuticals" and "Telecom Equipment" industries, indicating that the model may be valid to model dynamics in these industries. However, for the industries "Computers" and "Software", investment does not seem to adequately control for unobserved productivity differences so that the structural dynamic model has to be rejected for these industries.[18]

2.5.4 Robustness

Before we turn to the estimation of the conditional distribution of productivity, a note on robustness. We have conducted robustness checks with respect to the inclusion/exclusion of firms with a sales contribution of acquisitions exceeding 50%, the choice of depreciation rates in the construction of the capital stock, and the use of price deflators for investment and output (see Appendix A2.2). While the level of the point estimates differ slightly from case to case, the relative magnitude of the coefficients across specifications was relatively robust. Estimates were sensitive, however, to the method of imputation of wage and pension costs in the construction of value added, which is necessary as these items are missing for most companies in COMPUSTAT. As explained in the appendix, we experimented with a number of specifications, choosing the one that yielded a satisfactory fit between imputed and actual values.

[18] Rejection may also be due to a misspecification of the production function. However, moving from a Cobb-Douglas to a translog production function did not change the conclusions.

Is it worth investing in R&D?

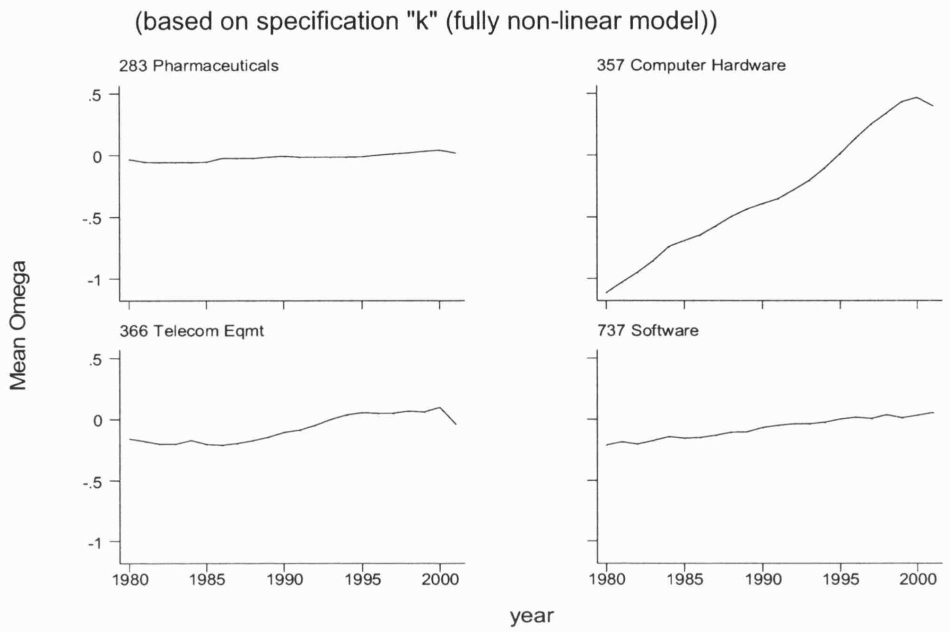
2.6 Effect of R&D on the distribution of productivity

The estimates of the firms' unobserved productivity state over time enable us to analyse the distribution of future productivity conditional on current productivity and R&D. This is an interesting exercise for a number of reasons: (1) Our model assumes that this conditional distribution is stochastically increasing in R&D. A rejection of this property would therefore clearly reject our model (at least in combination with a Cobb-Douglas production function). (2) In many models of firm dynamics, the driving force behind idiosyncratic outcomes paths across firms is a stochastic state variable (call it productivity) that follows a first order Markov process which may or may not be influenced by a the firm. The distribution of the future productivity state conditional on the current state and R&D is therefore the central primitive for dynamics in this and other models of firm dynamics. Our approach allows us to estimate the empirical counterpart of this primitive in the context of our model without having to solve the dynamic problem. (3) The literature on R&D and productivity has typically been concerned with estimating the average effect of R&D on productivity by putting R&D (or a deterministically accumulated stock of R&D) in the production function. Our approach explicitly treats the R&D process as one with stochastic accumulation and allows us to estimate the entire conditional distribution of productivity realisations. In that sense, it gives us a more complete picture of the effect of R&D on productivity.

value ?

In what follows, we use two alternative productivity estimates based on specifications "k" (i.e. the estimates produced by estimating the system of equations (2.7) and (2.18)) and "p" (the original Olley-Pakes specification based on estimating the system (2.7) and (2.17)) from the previous section. These specifications seem the most reasonable ones on the basis of the estimated production function coefficients, and because of the need to control for R&D and exit. We present results for all industries for both specifications despite the fact that the specification test in the previous section lead to the conclusion that specification "p" can be accepted for "Pharmaceuticals" and "Telecom Equipment" only and specification "k" for "Pharmaceuticals only. Note that in these specifications R&D data has

Figure 2-1: Weighted average of industry productivity over time



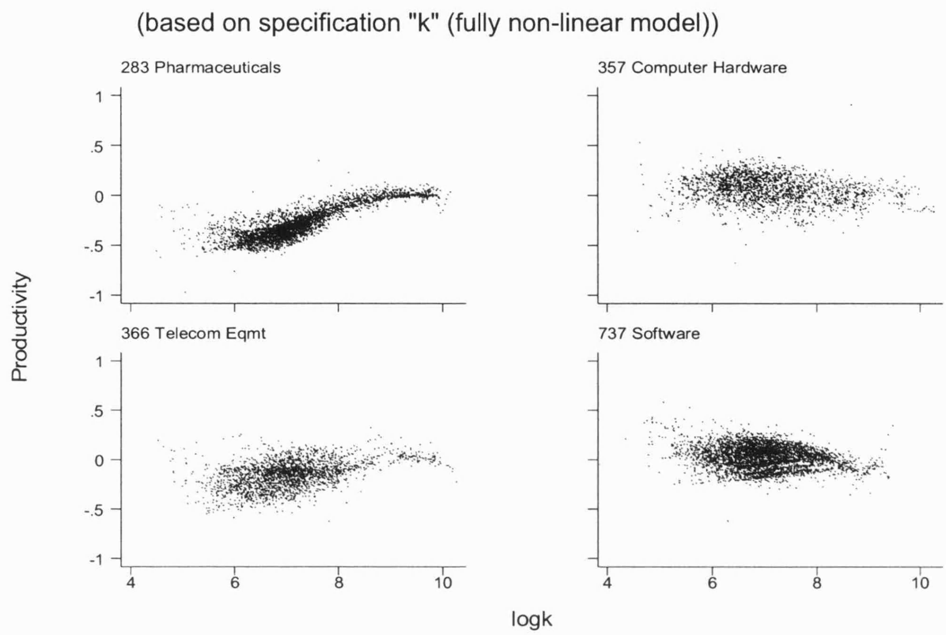
not entered the estimation of unobserved productivity. why do we care?

We start by graphically inspecting various features of the productivity estimates. Figure 2-1 shows the sales weighted average productivity over time across industries. The most striking feature is the huge increase in productivity in "Computer Hardware" over the past twenty years. The increase in productivity in this industry is a mirror image of the massive decline in quality adjusted output prices in this industry which we use to deflate value added. Productivity increases in the other industries are much more modest. For "Software", however, we do not put much faith in the aggregate productivity estimates, as we lack a specific output price deflator for this industry.

To remove any time effects or aggregate shocks from the analysis of the distribution of productivity, we use deviations from the sales weighted industry year average as productivity measure in our subsequent analysis. Figure 2-2 shows scatter plots of the (detrended) estimated productivity state against the capital stock for each industry. Note that only observations with positive investments are available, so that the lower bound on these scatter plots cannot be directly interpreted as the exit threshold $\underline{\omega}(k)$. Figure 2-3 shows histograms for the increment of productivity, $\Delta\omega_t = \omega_{t+1} - \omega_t$. The widest dispersion in these increments is observed

effect of business cycle?

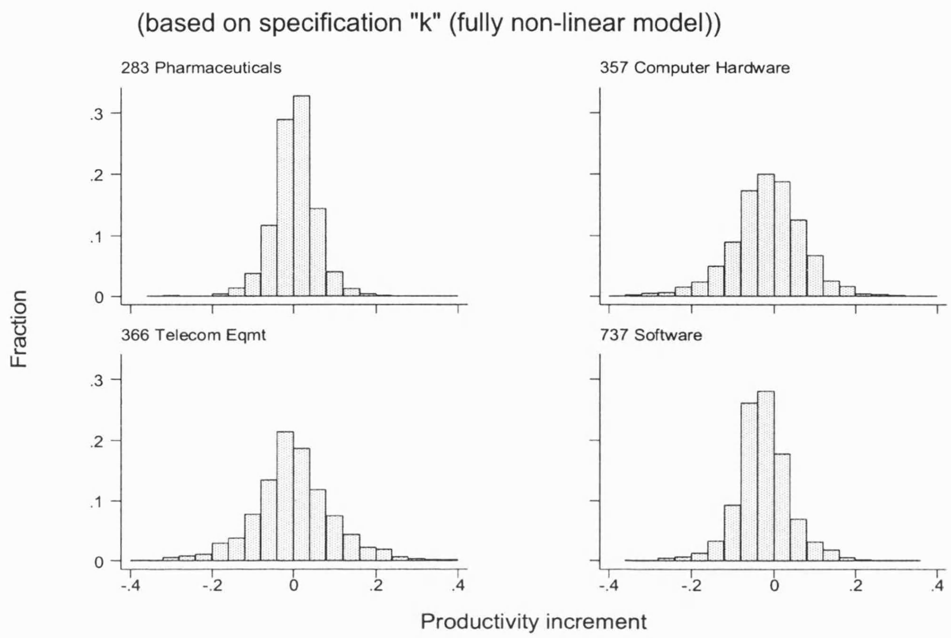
Figure 2-2: Productivity against capital stock by industry



for "Telecom Equipment" followed by "Computers", "Software", and "Pharmaceuticals". This hints towards at least quantitative differences in the dynamics of productivity across these industries.

Figure 2-4 is the main figure of interest. It plots empirical conditional future productivity distributions for productivity estimates based on specification "k". This allows us to graphically assess whether the empirical productivity distribution is stochastically increasing in R&D. To condition on current productivity and on R&D, we have partitioned the sample for each industry into observations with current productivity above/below the median productivity and into observations with R&D investments above/below the median R&D investment. Figure 2-4 plots the empirical cumulative distribution functions of the productivity increments $\Delta\omega_t$ by current productivity level (high/low) and current R&D spending (high/low) and by industry. For each industry, there are two graphs. The left graph corresponds to observations with low current productivity state and the right graph the observations with currently high levels of productivity. Each graph shows two empirical cdf's for productivity increments. The solid line is the cdf for observations with low R&D expenditure and the dashed line the cdf for observations with high levels of R&D investments.

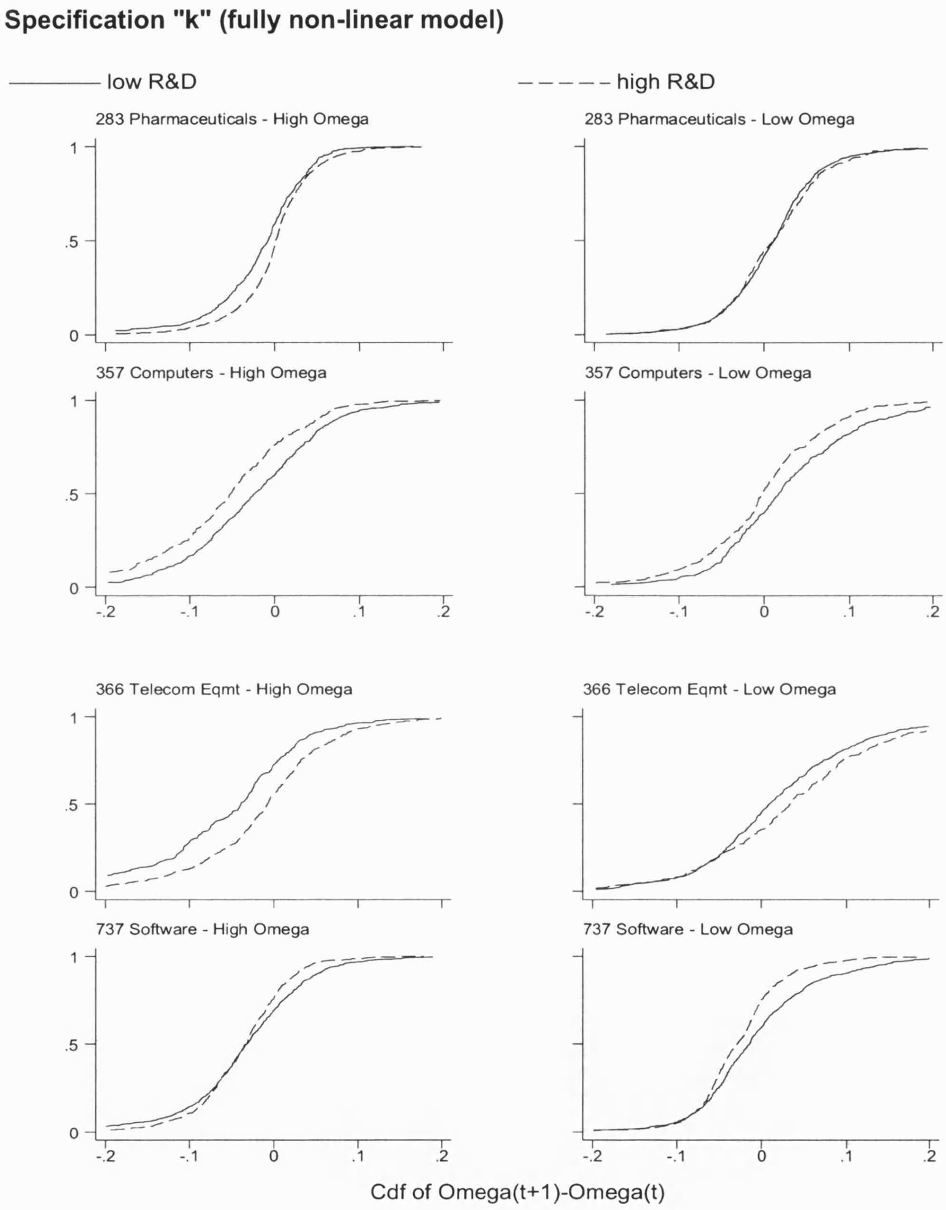
Figure 2-3: Histograms on productivity increments by industry



If the future productivity is in fact stochastically increasing in R&D, the dashed cdf for high spending firms should lie to the right of the solid cdf for low spending firms. The graphs seem to confirm the hypothesis of first order stochastic dominance of the productivity realisations in R&D for "Pharmaceuticals" and "Telecom Equipment", while the hypothesis seems to be rejected for "Computer Hardware" and "Software". So from this graphical representation, we would reject our dynamic model including R&D for "Computers" and "Software", while we would accept it for "Pharmaceuticals" and "Telecom Equipment". However, the graphs rely on discretising the continuous conditioning variables of current productivity and R&D.

Figure 2-5 shows the corresponding graphs for the productivity estimates produced by regression specification "p" – the original Olley-Pakes specification. The only apparent difference to Figure 2-4 is that for low productivity firms in the "Telecom Equipment" industry the cdfs for low and high spending firms almost coincide. If anything, the dashed cdf for firms spending significant amounts on R&D lies to the left of the cdf for low spenders. This is puzzling as, in contrast to specification "k" in Figure 2-4, specification "p" was accepted for this industry by the specification tests in the previous section.

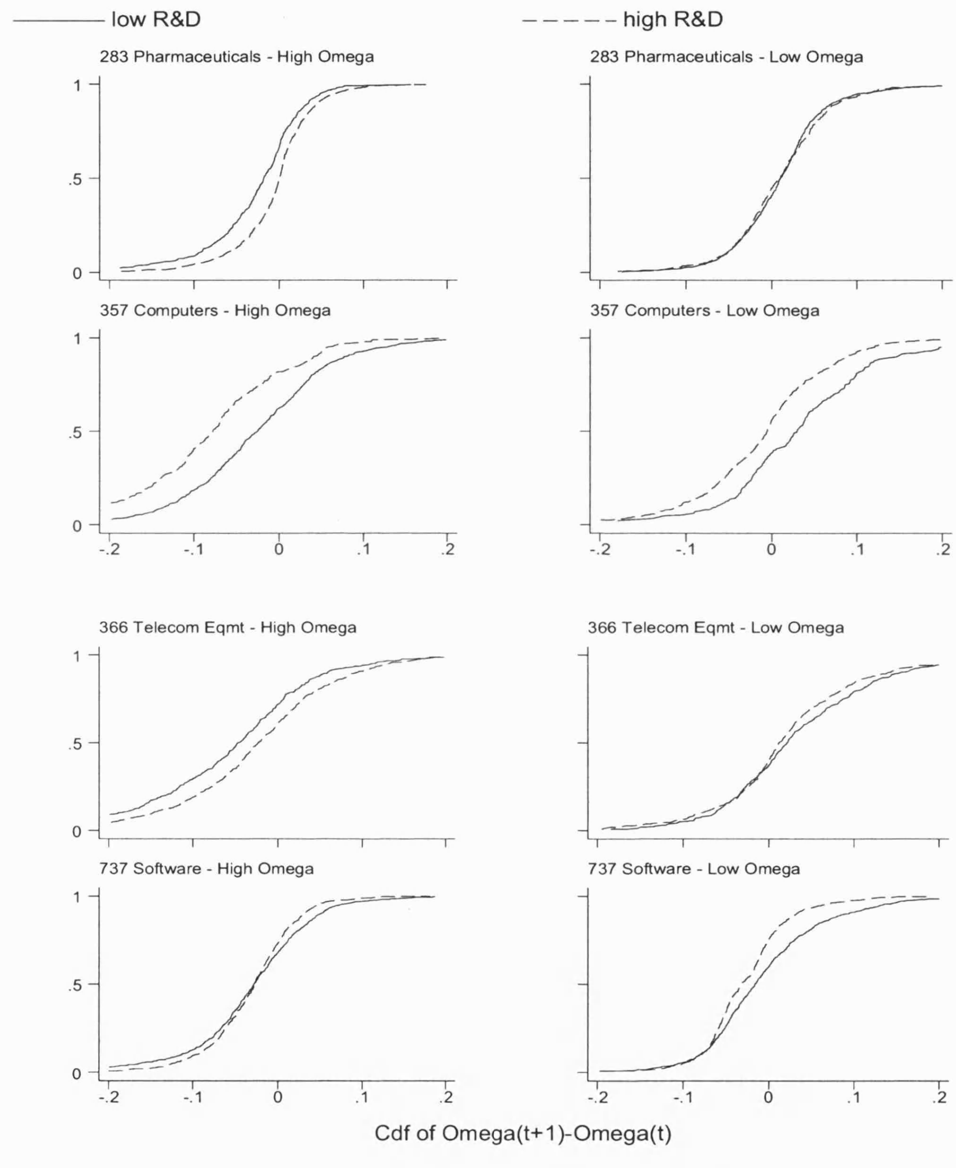
Figure 2-4: Distribution of productivity increments by industry, initial productivity level, and R&D spending



For each industry, the sample was partitionned into firms above/below the median productivity and firms with R&D investments above/below the medain R&D expenditure. The left panel for each industry shows the empirical cdf's for the productivity increments for low productivity firms and the right panel the cdf's for high productivity firms.
The solid line in each graph is the cdf for those firms spending relatively little on R&D and the dashed line the cdf for firms with high R&D spending.

Figure 2-5: Distribution of productivity increments by industry, initial productivity level, and R&D spending

Specification "p" (original Olley-Pakes specification)



For each industry, the sample was partitionned into firms above/below the median productivity and firms with R&D investments above/below the medain R&D expenditure. The left panel for each industry shows the empirical cdf's for the productivity increments for low productivity firms and the right panel the cdf's for high productivity firms.
The solid line in each graph is the cdf for those firms spending relatively little on R&D and the dashed line the cdf for firms with high R&D spending.

Table 2.9 presents simple regression results for specification "k" that confirm the conclusion from Figure 2-4. As in the figures, we do not control for censoring of the distribution through exit or through negative investments. Instead, we simply present the distributions conditional on survival and positive investments.

The regressions we run are OLS and quantile regressions of future productivity ω_{t+1} on current productivity ω_t and log R&D. While the OLS regressions estimate the mean effect of R&D (and current productivity) on productivity, the quantile regressions allow us to estimate the effect of R&D on different quantiles of the conditional distribution. We reject the hypothesis of first order stochastic dominance of future productivity in R&D if none of the coefficients on R&D is significantly positive or if at least one of the coefficients is significantly negative.

The regressions are clearly simplistic as they assume that there is a linear relationship between the conditional quantiles (respectively the mean) and the dependent variables. In the context of our model, this amounts to assuming that the choice ψ_t can be expressed as a linear function of R&D and current productivity and a linear relationship between ψ_t and the quantiles. However, these simple regressions serve as a first order approximations to quantify the effect of R&D on productivity.

For each industry, we also run a second set of regressions that include log capital. While this is not in line with our model which assumes that the distribution of productivity is only a function of current productivity and R&D, it allows us to check whether the stochastic dominance result is merely driven by the size of the firm. It may also help to control for the censoring problem induced by the exit threshold which is a function of the capital stock.

For "Pharmaceuticals" and "Telecom Equipment", the coefficient for R&D is positive in the OLS specifications in each of the quantile regressions irrespective of the quantile estimated. On the other hand, the coefficients on R&D are mostly negative for "Computers" and "Software" as we had suspected from the graphs. These patterns do not change once we also include log capital.

Below the point estimates we report standard errors of the coefficients on the basis of the 200 bootstrapped samples of the production function estimation. Note that this accounts for the fact that the productivity variables are estimated as the

Table 2.9: Quantile regressions on the conditional productivity distribution - specification "k"

Dependent variable: Omega(t+1) based on specification "k"
(fully non-linear specification)

a. Pharmaceuticals (SIC 283)

F.Omega	OLS	Quantile 0.25	Quantile 0.5	Quantile 0.75	OLS	Quantile 0.25	Quantile 0.5	Quantile 0.75
Omega	**0.838**	**0.920**	**0.886**	**0.858**	**0.745**	**0.826**	**0.845**	**0.828**
SE	0.098	0.088	0.070	0.070	0.087	0.087	0.080	0.088
Log_R&D	**0.021**	**0.017**	**0.014**	**0.012**	**0.006**	**0.003**	**0.008**	**0.008**
SE	0.010	0.009	0.007	0.009	0.006	0.007	0.006	0.007
Log_k					**0.030**	**0.028**	**0.012**	**0.010**
SE					0.018	0.014	0.014	0.018
const	**-0.188**	**-0.169**	**-0.124**	**-0.091**	**-0.330**	**-0.294**	**-0.179**	**-0.142**
SE	0.100	0.081	0.067	0.079	0.198	0.149	0.136	0.171
pseudo R^2	0.869	0.634	0.704	0.734	0.873	0.641	0.706	0.735
# obs	3023	3023	3023	3023	3023	3023	3023	3023

b. Computer Hardware (SIC 357)

F.Omega	OLS	Quantile 0.25	Quantile 0.5	Quantile 0.75	OLS	Quantile 0.25	Quantile 0.5	Quantile 0.75
Omega	**0.681**	**0.732**	**0.762**	**0.715**	**0.677**	**0.734**	**0.753**	**0.718**
SE	0.068	0.070	0.064	0.063	0.081	0.082	0.081	0.083
Log_R&D	**-0.016**	**-0.010**	**-0.012**	**-0.015**	**-0.008**	**-0.011**	**-0.005**	**0.003**
SE	0.012	0.013	0.012	0.013	0.013	0.014	0.013	0.013
Log_k					**-0.010**	**0.002**	**-0.008**	**-0.021**
SE					0.031	-0.001	-0.014	-0.029
const	**0.116**	**0.021**	**0.081**	**0.154**	**0.135**	**0.013**	**0.098**	**0.186**
SE	0.106	0.106	0.102	0.111	0.204	0.188	0.177	0.188
pseudo R^2	0.524	0.374	0.378	0.344	0.525	0.374	0.379	0.349
# obs	1821	1821	1821	1821	1821	1821	1821	1821

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

Dependent variable: Omega(t+1) based on specification "k"
(fuily non-linear specification)

c. Telecom Equlpment (SiC 366)

F.Omega	OLS	Quantile			OLS	Quantile		
		0.25	0.5	0.75		0.25	0.5	0.75
Omega	**0.620**	**0.680**	**0.696**	**0.631**	**0.616**	**0.685**	**0.689**	**0.620**
SE	0.050	0.070	0.053	0.053	0.063	0.080	0.063	0.057
Log_R&D	**0.031**	**0.034**	**0.026**	**0.023**	**0.040**	**0.031**	**0.037**	**0.045**
SE	0.017	0.016	0.015	0.019	0.013	0.014	0.012	0.013
Log_k					**-0.010**	**0.004**	**-0.013**	**-0.027**
SE					0.026	0.023	0.021	0.025
const	**-0.260**	**-0.324**	**-0.215**	**-0.152**	**-0.247**	**-0.327**	**-0.202**	**-0.110**
SE	0.153	0.136	0.135	0.170	0.208	0.170	0.168	0.212
pseudo R^2	0.540	0.376	0.372	0.337	0.541	0.376	0.373	0.344
# obs	1976	1976	1976	1976	1976	1976	1976	1976

d. Software (SIC 737)

F.Omega	OLS	Quantile			OLS	Quantile		
		0.25	0.5	0.75		0.25	0.5	0.75
Omega	**0.823**	**0.885**	**0.896**	**0.871**	**0.832**	**0.889**	**0.909**	**0.890**
SE	0.117	0.125	0.111	0.111	0.125	0.128	0.119	0.124
Log_R&D	**-0.008**	**0.001**	**-0.003**	**-0.013**	**-0.019**	**-0.017**	**-0.018**	**-0.023**
SE	0.008	0.009	0.008	0.009	0.010	0.012	0.009	0.010
Log_k					**0.015**	**0.025**	**0.018**	**0.012**
SE					0.014	0.016	0.013	0.015
const	**0.033**	**-0.064**	**-0.005**	**0.102**	**0.004**	**-0.120**	**-0.028**	**0.082**
SE	0.069	0.070	0.061	0.071	0.094	0.100	0.088	0.100
pseudo R^2	0.645	0.475	0.473	0.434	0.649	0.485	0.478	0.437
# obs	3089	3089	3089	3089	3089	3089	3089	3089

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

estimates will differ across bootstrap samples. As the coefficient estimates are often within two standard deviations of zero, it is hard to assess the significance of the coefficients from this table.

Table 2.10 reports the percentage of cases across the 201 samples in which the R&D coefficient exceeds zero. For "Pharmaceuticals" and "Telecom Equipment" very high percentages of the R&D coefficients are positive, so that we can reject the null of a nonpositive coefficient at conventional significance levels and conclude that the conditional productivity distribution is stochastically increasing in R&D for "Pharmaceuticals" and "Telecom Equipment". On the other hand, the table shows that we cannot reject the null of zero coefficients on R&D for "Computers" and "Software". This leads us to reject the hypothesis that the conditional productivity distribution is increasing in R&D for these two industries.

Tables 2.11 and 2.12 correspond to Tables 2.9 and 2.10 but use the productivity estimates from the original Olley-Pakes specification "p" in the regressions. Results are fairly similar with the exception of "Telecom Equipment". For this industry, as suggested by the comparison between Figures 2-4 and 2-5, the effect of R&D comes out much weaker with specification "p" than with "k". In fact, for the OLS and quantile regressions that do not include capital, the coefficients on R&D are virtually zero, even though the null hypothesis of zero coefficients can be rejected at a 10% significance level for the OLS coefficient and the 25% quantile and the median (Table 2.12, panel a). Once capital is included in the regression, the results are almost identical than those from specification "k". This is not surprising. After all the two specifications only differ with respect to the coefficient estimate for capital, i.e. the difference in the productivity estimates of the two specifications is proportional to capital.

Since we accept our dynamic model for "Pharmaceuticals" and "Telecom Equipment", we can in principle address the question of the rate of return to R&D implied by our model. Clearly, the long run expected returns to R&D would have to be measured by the marginal effect of R&D on the firm's value function. This measure would require a solution for the value function from the Bellman equation (2.1) which in turn would require the specification and estimation of all the primitives

Table 2.10: Significance of coefficient for log R&D - specification "k"

Specification "k" (fully non-linear specification)

a. Regressions of F.omega on omega and log_R&D

	SIC	OLS	Quantile 0.25	Quantile 0.5	Quantile 0.75
Pharmaceut.	**283**	0.985	0.988	0.980	0.980
Computers	**357**	0.575	0.627	0.604	0.545
Telecom Eqmt	**366**	0.970	0.973	0.965	0.933
Software	**737**	0.597	0.786	0.622	0.512

b. Regressions of F.omega on omega, log_R&D, and log_K

	SIC	OLS	Quantile 0.25	Quantile 0.5	Quantile 0.75
Pharmaceut.	**283**	0.940	0.881	0.968	0.983
Computers	**357**	0.664	0.604	0.704	0.821
Telecom Eqmt	**366**	1.000	0.990	1.000	1.000
Software	**737**	0.555	0.597	0.555	0.527

The number in each cell reports the percentage of cases across the 201 samples (one original sample + 200 bootstrapped samples) with a positive coefficient on R&D

Table 2.11: Quantile regressions on the conditional productivity distribution - specification "p"

Dependent variable: Omega(t+1) based on speclfication "p"
(orlglnal Olley-Pakes stage2 equation)

a. Pharmaceuticals (SiC 283)

F.Omega	OLS	Quantile 0.25	0.5	0.75	OLS	Quantile 0.25	0.5	0.75
Omega	**0.722**	**0.814**	**0.806**	**0.789**	**0.670**	**0.743**	**0.780**	**0.778**
SE	0.100	0.091	0.075	0.075	0.077	0.075	0.068	0.077
Log_R&D	**0.019**	**0.019**	**0.012**	**0.010**	**0.002**	**0.000**	**0.004**	**0.006**
SE	0.007	0.005	0.005	0.006	0.005	0.005	0.004	0.005
Log_k					**0.023**	**0.023**	**0.011**	**0.005**
SE					0.014	0.010	0.009	0.011
const	**-0.164**	**-0.180**	**-0.105**	**-0.062**	**-0.220**	**-0.225**	**-0.128**	**-0.076**
SE	0.066	0.047	0.041	0.056	0.156	0.102	0.092	0.112
pseudo R^2	0.694	0.483	0.552	0.551	0.705	0.496	0.555	0.552
# obs	3023	3023	3023	3023	3023	3023	3023	3023

b. Computer Hardware (SIC 357)

F.Omega	OLS	Quantile 0.25	0.5	0.75	OLS	Quantile 0.25	0.5	0.75
Omega	**0.681**	**0.717**	**0.740**	**0.718**	**0.609**	**0.650**	**0.679**	**0.633**
SE	0.068	0.056	0.053	0.066	0.081	0.092	0.087	0.086
Log_R&D	**-0.034**	**-0.026**	**-0.029**	**-0.032**	**-0.003**	**-0.005**	**-0.001**	**0.006**
SE	0.014	0.015	0.014	0.015	0.017	0.018	0.016	0.016
Log_k					**-0.046**	**-0.034**	**-0.040**	**-0.054**
SE					0.046	0.047	0.044	0.045
const	**0.279**	**0.168**	**0.229**	**0.306**	**0.422**	**0.288**	**0.348**	**0.466**
SE	0.119	0.128	0.123	0.131	0.302	0.320	0.311	0.324
pseudo R^2	0.702	0.515	0.509	0.464	0.716	0.521	0.520	0.483
# obs	1821	1821	1821	1821	1821	1821	1821	1821

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

Dependent variable: Omega(t+1) based on specification "p"
(original Oliey-Pakes stage2 equation)

c. Telecom Equipment (SiC 366)

F.Omega	OLS	Quantile			OLS	Quantile		
		0.25	0.5	0.75		0.25	0.5	0.75
Omega	**0.606**	**0.639**	**0.683**	**0.658**	**0.555**	**0.605**	**0.646**	**0.579**
SE	0.038	0.053	0.039	0.056	0.050	0.065	0.048	0.050
Log_R&D	**0.007**	**0.011**	**0.003**	**0.002**	**0.040**	**0.032**	**0.033**	**0.044**
SE	0.013	0.014	0.014	0.015	0.013	0.014	0.011	0.013
Log_k					**-0.038**	**-0.020**	**-0.033**	**-0.053**
SE					0.020	0.021	0.019	0.021
const	**-0.039**	**-0.120**	**-0.020**	**0.044**	**0.013**	**-0.113**	**0.017**	**0.141**
SE	0.117	0.122	0.117	0.134	0.146	0.144	0.140	0.168
pseudo R^2	0.383	0.246	0.258	0.245	0.397	0.249	0.268	0.263
# obs	1976	1976	1976	1976	1976	1976	1976	1976

d. Software (SIC 737)

F.Omega	OLS	Quantile			OLS	Quantile		
		0.25	0.5	0.75		0.25	0.5	0.75
Omega	**0.864**	**0.926**	**0.933**	**0.913**	**0.865**	**0.917**	**0.934**	**0.917**
SE	0.108	0.108	0.094	0.097	0.107	0.104	0.093	0.100
Log_R&D	**-0.004**	**0.003**	**-0.001**	**-0.010**	**-0.016**	**-0.013**	**-0.013**	**-0.020**
SE	0.006	0.007	0.005	0.005	0.009	0.012	0.009	0.010
Log_k					**0.015**	**0.022**	**0.015**	**0.013**
SE					0.009	0.011	0.009	0.010
const	**0.002**	**-0.085**	**-0.021**	**0.078**	**-0.025**	**-0.130**	**-0.046**	**0.056**
SE	0.037	0.044	0.032	0.033	0.043	0.050	0.039	0.039
pseudo R^2	0.663	0.498	0.490	0.436	0.667	0.506	0.494	0.439
# obs	3089	3089	3089	3089	3089	3089	3089	3089

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

Table 2.12: Significance of coefficient for log R&D - specification "p"

Specification "p" (orig. Olley-Pakes)

a. Regressions of F.omega on omega and log_R&D

	SIC	OLS	Quantile		
			0.25	0.5	0.75
Pharmaceut.	283	1.000	1.000	1.000	0.993
Computers	357	0.520	0.565	0.537	0.512
Telecom Eqmt	366	0.935	0.965	0.898	0.853
Software	737	0.662	0.888	0.709	0.510

b. Regressions of F.omega on omega, log_R&D, and log_K

	SIC	OLS	Quantile		
			0.25	0.5	0.75
Pharmaceut.	283	0.911	0.829	0.958	0.973
Computers	357	0.697	0.684	0.731	0.843
Telecom Eqmt	366	1.000	0.988	1.000	1.000
Software	737	0.572	0.612	0.575	0.532

The number in each cell reports the percentage of cases across the 201 samples (one original sample + 200 bootstrapped samples) with a positive coefficient on R&D

of the model, which is beyond the scope of this study.

However, the OLS coefficients on R&D from Table 2.9 and 2.11 give a crude estimate of the effect of R&D on expected productivity. Since the coefficient on productivity in the production function (2.5) is normalised to unity and ignoring the fact that $E(\log y) \neq \log E(y)$, the OLS coefficients in the first column of tables 2.9 and 2.11 are a rough approximation of the elasticity of expected value added in the next period with respect to R&D. So the point estimate of this elasticity is about .02 for "Pharmaceuticals" and, depending on the specification ranges between .007 and .04 for "Telecom. Equipment". Both estimates appear rather low.[19] However, they are only a measure of the short run returns to R&D (and a crude one at that). If we assume that the one period "shift" in the distribution of productivity is in fact permanent the long run return to R&D in terms of value added would be the discounted value of the a permanent increase in value added. For a discount rate of 10% this would amount to a long run elasticity of value added with respect to R&D of approximately .2 for "Pharmaceuticals" and up to .4 for "Telecom. Equipment".

[19] In traditional firm level studies of R&D on productivity, estimates of the elasticity of output with respect to R&D that are based on a within or time-series approach (rather than cross sectional) are typically rather low (Mairesse & Sassenou (1991)).

Presentation of results are so much improved!

2.7 Conclusion

This study develops an empirical framework for the analysis of the effect of R&D on the entire distribution of a firm's future productivity conditional on its current productivity. Its basis is a structural model for firm dynamics in which firm invest in physical capital and in R&D. We prove that the investment policy function in this model is monotonic in productivity (conditional on capital), which allows us to employ an approach similar to Olley & Pakes (1996) to jointly estimate production function coefficients and firms' unobserved productivity state over time in the context of R&D. We use these estimates to analyse the conditional distribution of productivity. The main conclusions from this study are as follows.

First, the invertibility result implies that investment is a valid proxy for productivity in the model with R&D and capital investments. Therefore, an estimation approach similar to Olley & Pakes (1996) can be applied to control for unobservable productivity differences across firms to estimate production function coefficients of variable factors even if the model includes investments in R&D and in physical capital.

Second, the original stage two estimation equation of the Olley-Pakes procedure for the estimation of quasi-fixed factors of production is shown to be valid in this model and asymptotically equivalent to the alternatives developed in this study. Therefore, both stages of the Olley-Pakes procedure are valid in our model. However, the analysis shows that this relies on a specific property of the model. In general, neglecting R&D in stage two of the Olley-Pakes estimation procedure may lead to inconsistent estimates of the coefficients of quasi-fixed factors of production. Whenever there is a link between R&D and future productivity, one needs to be explicit about the underlying dynamic model for R&D and analyse, how to adequately control for the expected productivity in stage two of the estimation. This applies independently of the model in this chapter and regardless of whether one uses investment or intermediate inputs to proxy for productivity.

Third, estimates of the conditional productivity distribution can be used to test our model. They also provide a basis for analysing the effect of R&D on the entire distribution of productivity. This sheds light on the key primitive of many theoretical models of firm dynamics and allows us not only to estimate the average short run returns to R&D but to analyse the entire distribution of these returns.

Unfortunately, answering questions of long run returns to R&D and, more generally, analysing the implication of policy changes on firm dynamics and investment behaviour more conclusively still requires a solution to the full dynamic model. This is precisely what the estimation approach here avoids. Its virtue is that it allows us to estimate a subset of the primitives of the dynamic model in a relatively simple and straightforward way without having to estimate the full dynamic model. Estimation of the full model involves specifying functional forms of all the model primitives (not just the production function) and optimising over the full vector of parameters using a nested fixed point approach that solves the dynamic model for each new value of the parameters. This is beyond the scope of this study but would allow us to answer a range of questions more precisely.

Chapter 3

Productivity Dynamics in the Presence of Unobserved Knowledge Capital

3.1 Introduction

This chapter extends the model and the empirical techniques of the previous chapter to include an unobserved knowledge capital state. Knowledge capital allows for lagged effects of R&D on productivity and is modelled as the result of a partly deterministic and partly stochastic accumulation process. This integrates the traditional Griliches approach of deterministic accumulation of R&D into knowledge capital with the Ericson & Pakes (1995) approach of stochastic accumulation of R&D into productivity. Methodologically, the introduction of knowledge capital requires us to extend the original invertibility approach by Olley & Pakes (1996) for one unobserved state variable to our model with two unobserved states variables. We show that, under certain assumptions, observed R&D investments can serve as an additional proxy for the second state variable in the estimation. The empirical focus lies on assessing how the introduction of unobserved knowledge affects the production function estimates and the effect of R&D on future productivity. The elasticity of productivity with respect to knowledge capital also allows us to analyse to importance of lagged effects of R&D on productivity.

The model in the previous chapter takes an extreme approach as to how R&D accumulates into productivity. Conditional on the current productivity state, R&D investments improve the distribution of the next period's productivity state (in the FOSD sense) and hence the next period's payoff. The effect of R&D indirectly also transmits into more distant periods through serial correlation in productivity, but there is no direct effect of current R&D on productivity realisations beyond the next period. This stochastic accumulation process for the productivity state is in line with the theoretical literature on firm dynamics (e.g. Ericson & Pakes (1995)) but raises the question of possible time lags between the R&D investments and the point in time when the outcome of the R&D process becomes payoff relevant.

At the other extreme, there is the view that R&D accumulates deterministically into knowledge capital in the tradition of Griliches (e.g. Griliches (1998) for a collection of papers in this tradition). In this literature knowledge capital becomes an input in the production function and there is no inherent difference in the accumulation of knowledge capital and that of physical capital. Both forms of capital accumulate through investments and depreciate in a deterministic way. This approach can deal with lagged effects of R&D on production to the extent that knowledge depreciates slowly. However, the researcher has to specify a depreciation rate for knowledge to construct the knowledge stock from past R&D investments, which is to some extent arbitrary. Furthermore, the deterministic accumulation approach neglects the inherently stochastic nature of the R&D process.

To reconcile these views, we introduce a knowledge state into the model of the previous chapter. Rather than being a direct input into the production function, the role of knowledge is to improve the distribution of future productivity (in the first order stochastic dominance (FOSD) sense). The firm's knowledge state is the result of an accumulation process containing stochastic as well as deterministic elements. Knowledge in the next period depends deterministically on the current knowledge state, on R&D investments, and on the stochastic realisation of productivity. The next period's knowledge state then determines the distribution of next period's productivity.

The combination of deterministic and stochastic elements in the accumulation process offers the advantage to allow for lags in the direct effects of R&D of more than one period while at the same time accounting for the stochastic nature of R&D. In this way, the approach can integrate the advantages of the traditional Griliches view of deterministic accumulation of knowledge through R&D with that of the stochastic accumulation approach of the theoretical firm dynamics literature.

A direct implication of the role of stochastic realisations of unobserved productivity in the accumulation process is that knowledge itself becomes an unobserved state variable in the model. This complicates the theoretical and empirical analysis, as it rules out the construction of knowledge capital from past R&D investments. Therefore, we require techniques that do not rely on specifying a depreciation rate for knowledge to construct knowledge capital. When feasible, the lack of dependence on a depreciation rate for knowledge constitutes an additional advantage over the Griliches approach as this depreciation rate typically has to be chosen in an ad hoc manner.

Moreover, the fact that knowledge and productivity are both unobserved state variables in the model raises an interesting methodological issue as to how to control for these states. Olley & Pakes (1996) propose an invertibility approach in the context of a single unobserved state variable. In this chapter, we will explore whether this invertibility idea can still be applied when there are two unobserved states in the model. As emphasised by Olley & Pakes (1996) and in the previous chapter, the feasibility of the invertibility approach depends on the precise structure of the underlying model and needs to be examined on a case by case basis. This structure will be even more critical in the presence of two unobserved state variables. Therefore, this chapter can by no means offer a general answer to this question.

However, we show that under certain additional assumptions, the invertibility approach can still be applied to the model of this chapter. Under these assumptions, the policy functions for investment and R&D can jointly be inverted to yield unobserved productivity as a unique function of the observed capital state and observed investments in R&D and physical capital. This is a powerful theoretical result which provides a basis for controlling for unobserved state variables in empirical applications.

One example of an empirical application using invertibility results is the semiparametric approach by Olley & Pakes (1996) for the estimation of production functions in the presence of unobserved serially correlated productivity differences across firms. We extend this approach to the model with two unobserved states and R&D by including R&D in the list of productivity proxies. Estimation proceeds in two stages. While conditions underlying the invertibility results are sufficient to identify the coefficients of the variable factors of production in the first stage, an additional condition is required for R&D to be a valid proxy in the second stage of the estimation which estimates the coefficients of quasi-fixed factors of production. Unfortunately, this is a condition on the second derivative of the expected future value of the firm, i.e. not a condition on the model primitives but the solution of the dynamic model. Therefore, this condition cannot be checked without specifying all the model primitives and solving the dynamic programme – the difficult task the semiparametric approach aims to avoid in the first place. Therefore, if one wants to follow the semiparametric approach, as we do, one can only hope that this condition is satisfied and proceed with the estimation.

Using the COMPUSTAT dataset of the previous chapter for the 3-digit SIC industries "Pharmaceuticals (SIC 283)", "Computer Hardware (SIC 357)", "Telecommunications Equipment (SIC 366)", and "Software (SIC 737)", we find that the production function estimates for the model with unobserved knowledge change somewhat, but not dramatically compared to the estimates of the model without knowledge state. The labour coefficients in the model with unobserved knowledge tend to be a little higher than that of the previous chapter, while the estimated capital coefficients tend be a little lower. This indicates that if the true model includes knowledge capital, the OLS biases in the labour coefficient due to simultaneity of the input choices may be quantitatively somewhat less important than suggested by the results of the previous chapter, even though they are still present.

Using the same specification tests as in the previous chapter, we find that the model with knowledge capital is rejected for all industries except "Pharmaceuticals". In that sense, introducing knowledge in the model and including R&D in the list of proxies does not improve on the model rejections in the previous chapter. In fact, for "Telecom Equipment" we accepted one of the estimation specifications in the previous chapter while we reject both knowledge capital specifications here. The rejections suggest that for these industries investment and R&D are insufficient to proxy for unobserved differences in productivity and knowledge.

Nevertheless, it is interesting to examine the effect of R&D and knowledge on productivity dynamics. We find that the estimates for the short run elasticity of productivity with respect to R&D is around .016 in Pharmaceuticals and around .03 in "Telecom Equipment" which is quite close to the estimates of the previous chapter. We also find that the elasticity of productivity with respect to knowledge is around .4 in "Pharmaceuticals", while it is much lower in the other industries. As knowledge capital allows for lagged effects of R&D on productivity, this confirms the prior that there are typically long lags in the pharmaceutical industry between the innovation of a new drug and bringing it to the market. It also suggests that introducing unobserved knowledge capital is an important improvement for modelling this industry.

The chapter is organised as follows. Section 3.2 introduces the model with unobserved knowledge capital and develops the theoretical invertibility result. Section 3.3 extends the estimation approach from Olley & Pakes (1996) and the previous chapter to the model with two unobserved states and derives conditions under which R&D is a valid proxy for the second stage of the estimation. Section 3.4 presents results of the estimation of production function coefficients. Section 3.5 analyses the effect of R&D and unobserved knowledge on the future productivity distribution. Section 3.6 concludes.

3.2 The theoretical model

3.2.1 Structure of the model

The model is an augmented version of the stochastic dynamic single agent model of the previous chapter which includes a knowledge state ψ. Firms are assumed to maximise the expected discounted value of future net cash flows and have now three state variables: productivity ω, capital k and knowledge ψ. At the beginning of each period, a firm observes its state and makes a discrete decision whether to exit or continue in operation. If it exits, it receives a termination value Φ. Otherwise, it earns current period profits $\pi(\omega, k)$ and decides how much to invest in physical capital and in R&D.

We introduce the unobserved knowledge state in the model by reinterpreting the choice of distribution ψ' of the previous chapter as the firm's knowledge state for the next period. This knowledge state ψ' determines the distribution of next period's productivity state ω'. This distribution is a member of the family of distributions

$$\mathcal{F}_{\psi'} = \{F(\omega'|\psi'), \psi' \in \Psi\}.$$

Distributions in this family are stochastically increasing in the next period's knowledge state ψ' in the first order stochastic dominance sense. As before, imposing a complete ordering of all the members of $\mathcal{F}_{\psi'}$ by indexing them by a single variable ψ' is a severe restriction which is crucial to derive the monotonicity results we rely on.

The knowledge state accumulates partly stochastically and partly deterministically in the tradition of Griliches. That is, the next period's knowledge state ψ' depends in a deterministic way on the current knowledge state ψ, the current amount of R&D investments, and the current stochastic productivity realisation ω. Note that if next period's knowledge state ψ' only depended on R&D and the inherited stock ψ but not on the stochastic productivity realisation, the model would reduce to a purely deterministic accumulation of knowledge capital in the tradition of Griliches. Similarly, if the knowledge ψ' did not depend on the inherited knowledge stock ψ but on R&D and the current productivity state, the model would reduce to the stochastic accumulation model of the previous chapter. When R&D, current productivity ω, and the current knowledge state ψ jointly determine the next period's knowledge state ψ' the accumulation process is a mixture of a deterministic and a stochastic accumulation.

As in the previous chapter, the key to modelling this accumulation process lies in specifying an R&D function, which specifies the amount of R&D the firm needs to spend to reach a given knowledge state ψ', conditional on the productivity realisation ω and the inherited stock ψ. This R&D function is denoted $r(\psi', \omega, \psi)$.

It is increasing in ψ' as improving knowledge is costly and decreasing in the current knowledge ψ and the productivity realisation ω as firms with higher endowments require less R&D to reach a given knowledge state ψ'. The difference to the model in the previous chapter is that this function includes the current knowledge state ψ. To obtain the invertibility results, the function needs to satisfy a number of assumptions which are detailed and discussed below. However, up to these assumptions, the function can be left unspecified. In particular, we do not need to specify an arbitrary depreciation rate for knowledge as in the Griliches type literature.

As before, the deterministic accumulation of physical capital k follows the usual equation

$$k' = (1-\delta)k + i,$$

where k' denotes next period's capital stock, δ the rate of capital depreciation and i the investment choice of the firm. The cost of investment are denoted $c(k', k)$.

With these specifications in hand, the Bellman equation of the modified dynamic model with discount factor β becomes:

$$V(\omega, k, \psi) = \max\left\{\Phi, \sup_{k',\psi'}\left[\pi(\omega, k) - c(k', k) - r(\psi', \omega, \psi) + \beta \int V(\omega', k', \psi') dF(\omega'|\psi')\right]\right\}. \tag{3.1}$$

The only new element in this equation is the inclusion ψ in the value function $V(\omega, k, \psi)$ and in the R&D function $r(\psi', \omega, \psi)$. The firm's controls are a discrete exit decision, and continuous choices of next period's capital stock k' and knowledge stock ψ'. The maximum expected discounted value of a firm $V(\omega, k, \psi)$ is the larger of the sell-off value Φ and expected discounted value of continuation. The latter consists of current period profits $\pi(\omega, k)$ reduced by the cost of investment $c(k', k)$ and by R&D expenditure $r(\psi', \omega, \psi)$ plus the expected discounted value from the next period onwards.

The policy functions for exit, capital and knowledge depend on the full vector of state variables which now includes the knowledge state ψ:

$$\text{Exit rule} \quad : \quad \chi = \begin{cases} 1 & \text{(continue) if } \omega \geq \underline{\omega}(k, \psi) \\ 0 & \text{(exit) otherwise} \end{cases} \tag{3.2}$$

$$\text{Capital choice} \quad : \quad k' = \widetilde{k}(\omega, k, \psi) \tag{3.3}$$

$$\text{Knowledge choice} \quad : \quad \psi' = \widetilde{\psi}(\omega, k, \psi) \tag{3.4}$$

The particular form of the exit rule in equation (3.2) results from the fact that the profit function is increasing and the R&D function decreasing in ω which implies that the continuation value must be increasing in current productivity ω.

3.2.2 Invertibility with two unobserved state variables

The invertibility argument with the two unobserved state variables productivity and knowledge involves two steps. The first is to show the monotonicity of the policy functions (equations (3.3) and (3.4)) in the state variables. The argument is similar to the one in the previous chapter and requires similar assumptions on the extended R&D function.

However, monotonicity of the policy functions will merely define a schedule of combinations of productivity and knowledge that is consistent with the observed capital states in the current and the next period. The second step in the theoretical argument involves deriving additional conditions on the R&D function that ensure that observed R&D investments are sufficient to pin down productivity and knowledge. Under these conditions, unobserved productivity and knowledge can be expressed as functions of observed capital, observed investments in physical capital, and observed R&D investments.

The assumption on the profit function $\pi(\omega, k)$, the cost of physical capital investment $c(k', k)$, and the distribution function $F(\omega'|\psi')$ are the same as in the previous chapter (Assumptions A0 to A4) and will not be repeated here. The R&D function $r(\psi', \omega, \psi)$ now has the current knowledge state ψ as an additional argument and the assumptions on this function are augmented as follows.

Assumption *The R&D function* $r(\psi', \omega, \psi)$ *is*

(A1.c') *nonnegative,*

(A2.c') *increasing in the next period's knowledge state* ψ', *and decreasing in productivity* ω *and knowledge* ψ,

(A3.c') *submodular in* (ψ', ω, ψ), *and*

(A4.c') *twice continuously differentiable with* $\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \omega} < 0$ *and* $\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \psi} < 0$ *on the set* $\{(\psi', \omega, \psi) | r(\psi', \omega, \psi) > 0\}$.

As before, we first establish supermodularity and other properties of the value function in Lemma 3.1 which will be crucial for the monotonicity results.

Lemma 3.1 *The* ***value function*** $V(\omega, k, \psi)$ *of the model in equation (3.1) is*

1. *bounded under (A0),(A1),*
2. *nondecreasing in productivity* ω, *capital* k, *and knowledge* ψ *under (A0),(A2),*
3. *unique under (A0),(A1), and*
4. *supermodular under (A0) to (A3).*

Proof. *The proofs of boundedness, monotonicity and uniqueness of the value function follow the proof of Lemma 2.1 in the previous chapter. Supermodularity hinges on whether the expectation in equation (3.1) can be shown to be supermodular. Topkis (1998, Theorem 3.10.1) shows that* $\int V(\omega', k', \psi') dF(\omega'|\psi')$ *is supermodular in* (k', ψ') *if* $F(\omega'|\psi')$ *is stochastically increasing in* ψ' *and provided is* $V(\omega', k', \psi')$ *is supermodular in* (ω', k', ψ'). *As in the previous chapter, supermodularity then follows as the limit of an induction proof for a finite horizon problem.* ■

Lemma 3.2 corresponds to Theorem 2.4 in the previous chapter and establishes the monotonicity of the policy functions.

Lemma 3.2 *The policy functions for the* ***capital*** $\widetilde{k}(\omega, k, \psi)$ ***and knowledge*** $\widetilde{\psi}(\omega, k, \psi)$ *are*

1. *nondecreasing in ω, k, and ψ under (A0) to (A3), and*

2. *strictly increasing in ω on the set $\{(\omega, k, \psi)|\widetilde{k}(\omega, k, \psi) > k \wedge r(\widetilde{\psi}(\omega, k, \psi), \omega, \psi) > 0\}$ under (A0) to (A4).*

Proof. *1. Nondecreasing: Follows from Topkis (1978, Theorem 6.1) as the supermodularity assumptions, Lemma 3.1 and Topkis (1998, Theorem 3.10.1) imply that the objective function in the Bellman equation (3.1) is supermodular.*

2. Strictly increasing: Follows the proofs in Lemma 2.3 and Theorem 2.4 in the previous chapter. ■

Strict monotonicity of the capital policy function $\widetilde{k}(\omega, k, \psi)$ in productivity ω in Lemma 3.2 implies that there is a downward sloping schedule in (ω, ψ) space of combinations of productivity and knowledge that is consistent with observed values of k' and k. To pin down unique values for productivity and knowledge that are consistent with the data an additional piece of information is required. Theorem 3.3 establishes additional conditions on the R&D function that ensure that observed R&D investments provide this information.

Theorem 3.3 *Assume that assumptions (A0) to (A4) hold (including the modifications in this section), that the value function is twice continuously differentiable, and that the R&D function satisfies one of the additional conditions*

$$\frac{\partial r}{\partial \omega}\frac{\partial^2 r}{\partial \psi' \partial \psi} < \frac{\partial r}{\partial \psi}\frac{\partial^2 r}{\partial \psi' \partial \omega} \quad or \quad \frac{\partial r}{\partial \omega}\frac{\partial^2 r}{\partial \psi' \partial \psi} > \frac{\partial r}{\partial \psi}\frac{\partial^2 r}{\partial \psi' \partial \omega}.$$

everywhere.

Then the policy functions and the R&D function can be inverted to express productivity ω and knowledge ψ as unique functions of the observed capital stock k, the capital choice k', and the observed R&D investments r,

$$\begin{aligned} \omega &= \overline{\omega}(k', k, r) \\ \psi &= \overline{\psi}(k', k, r). \end{aligned}$$

Proof. *Conditional on k' and k, the investment policy function (3.3) implicitly defines a function in (ω, ψ) space of combinations of productivity and current* *knowledge states that are consistent with the observed k' and k. By the result of Lemma 3.2 this function is decreasing. The proof proceeds by showing that the conditions on the R&D function are sufficient for R&D to be strictly monotonically increasing or strictly monotonically decreasing along this function. Either of this is sufficient for (k, k', r) to uniquely define ω and ψ.*

First note that conditioning on k' and k also pins down the unobserved next period's knowledge state ψ'. This is because the optimal capital choice k' can be expressed as a function of ψ' and k from the Bellman equation (3.1) and is strictly increasing in ψ' using the results of Lemma 3.2. Therefore, ψ' is constant along the function in (ω, ψ) space defined by the investment policy function conditional on k' and k. This implies that the first order condition for choosing the next period's knowledge state has to hold along this function:

$$FOC = -\frac{\partial r(\psi', \omega, \psi)}{\partial \psi'} + \frac{\partial}{\partial \psi'}\beta E(V(\omega', k', \psi')|\psi') = 0$$

The slope of the function then follows from taking the differential of the FOC with respect to ψ and ω and setting it to zero,

$$\begin{aligned} dFOC &= -\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \psi} d\psi - \frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \omega} d\omega = 0, \text{ or} \\ \frac{d\omega}{d\psi} &= -\frac{\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \psi}}{\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \omega}} < 0. \end{aligned}$$

With the slope of the function in (ω, ψ) space in hand, we can now investigate whether observed R&D investments are monotonic along this function by forming the total differential of the R&D function holding ψ' constant.

$$\begin{aligned} dr &= \frac{\partial r(\psi', \omega, \psi)}{\partial \omega} d\omega + \frac{\partial r(\psi', \omega, \psi)}{\partial \psi} d\psi \\ &= \left[-\frac{\partial r(\psi', \omega, \psi)}{\partial \omega} \frac{\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \psi}}{\frac{\partial^2 r(\psi', \omega, \psi)}{\partial \psi' \partial \omega}} + \frac{\partial r(\psi', \omega, \psi)}{\partial \psi} \right] d\psi \end{aligned}$$

So R&D is strictly monotonically increasing (decreasing) in ω for fixed k', k if the term in brackets is strictly negative (positive) everywhere which is the condition *stated in the theorem. This completes the proof.* ■

Theorem 3.3 provides the empirically powerful invertibility result for the case of two unobserved state variables, productivity and knowledge. The burden of this result in terms of additional conditions on the model primitives rests entirely on the R&D function.[1] R&D functions that satisfy the monotonicity and submodularity assumptions (A1.c') to (A4.c') are, of course, not guaranteed to also satisfy the assumptions of Theorem 3.3. Whether the additional conditions in the theorem are likely to be satisfied in practice is an issue which will not be explored in any detail here.[2] However, it is easy to find R&D functions that satisfy all the necessary conditions at least for some parameter values. The following example illustrates this point.

Example 3.4 *Let the R&D function take the form*

$$r(\psi', \omega, \psi) = a\psi' - b\omega - c\psi - d\psi'\omega - e\psi'\psi - f\omega\psi$$

where $a, b, c, d, e, f > 0$ and where the domains of ψ', ω, ψ are normalised to be non-negative and bounded from above. These normalisations ensure that the sub-modularity assumptions are satisfied and that $\frac{\partial r}{\partial \omega} < 0$, and $\frac{\partial r}{\partial \psi} < 0$ as required, while $\frac{\partial r}{\partial \psi'} = a - d\omega - e\psi > 0$ will hold provided the parameter a is sufficiently large. Define

$$\begin{aligned} A &\equiv \frac{\partial r}{\partial \psi}\frac{\partial^2 r}{\partial \psi' \partial \omega} = (c + e\psi' + f\omega)d \\ B &\equiv \frac{\partial r}{\partial \omega}\frac{\partial^2 r}{\partial \psi' \partial \psi} = (b + d\psi' + f\psi)e. \end{aligned}$$

The condition of Theorem 3.3 then requires that either $A > B$ or $B > A$. This

[1] Note that we have assumed above, that $\frac{\partial^2 r}{\partial \psi' \partial \psi} < 0$ on the relevant set, while for the preceding results $\frac{\partial^2 r}{\partial \psi' \partial \psi} \leq 0$ would be sufficient. In fact, the additional conditions of Theorem 3.3 are trivially satisfied if $\frac{\partial^2 r}{\partial \psi' \partial \psi} = 0$. However, with $\frac{\partial^2 r}{\partial \psi' \partial \psi} = 0$ the inherited knowledge stock ψ would merely affect the level of R&D investments, not the choice of the next period's knowledge stock. The future knowledge ω stock would then be a function of k' and k only and observed R&D would not be necessary for the inversion.

[2] Checking whether the R&D function satisfies the additional conditions ex post is difficult as all three arguments are unobserved, while the estimation procedure only produces direct estimates of the productivity state ω.

can easily be satisfied by further constraining the parameters. E.g. making c (respectively b) sufficiently large ensures that $A > B$ (resp. $B > A$).

3.3 Estimation approach

The estimation approach follows closely the semiparametric estimation strategy for production functions developed in Olley & Pakes (1996) and will only be discussed briefly. The semiparametric estimation method proceeds in two stages. While the first stage of the follows directly from the theoretical results above, an additional difficulty arises in the second stage of the estimation procedure with two unobserved state variables. To resolve this problem, further additional conditions on the R&D functions are required.

Stage one: Estimation of the coefficients of the variable input(s)

Theorem 3.3 implies that unobserved productivity can be expressed as a function of the current and future capital stock and observed R&D investments so that the (Cobb-Douglas) production function can be rewritten as

$$\begin{aligned} y &= \alpha_0 + \alpha_l l + \alpha_k k + \omega + \eta \\ &= \alpha_l l + \phi(k', k, r) + \eta, \text{ where} \qquad (3.5) \\ \phi &= \phi(k', k, r) \equiv \alpha_0 + \alpha_k k + \overline{\omega}(k', k, r). \end{aligned}$$

Estimation of this equation as a partially linear semiparametric model (Robinson 1988) yields an estimate of α_l and estimates of the unknown function $\phi()$. Note that for the estimation of α_l, some variation in l that is uncorrelated with the state variables and hence to k', k, and r is required.

Stage two: Estimation of the coefficients of the quasi-fixed input(s)

For the second stage of the estimation procedure, rearrange the production function (3.5) to define a transformed dependent variable y^*

$$y^{*\prime} \equiv y' - \alpha_l l' = \alpha_0 + \alpha_k k' + \omega' + \eta',$$

where primes denote the values of variables in the next period as in the theoretical section. Because of the Markov assumption for the productivity process, this equation can be rewritten as

$$y^{*\prime} = \alpha_0 + \alpha_k k' + E[\omega'|\psi', \chi' = 1] + \xi' + \eta', \tag{3.6}$$

where the productivity innovation ξ' is uncorrelated with k' and ψ'. To estimate α_k consistently from this equation, we have to control for the expected productivity conditional on survival. The problem here lies in controlling for ψ' without losing the identification of α_k. Note that α_k enters not only the linear term in equation (3.6), but also ϕ from stage 1. Conditional on α_k, $\phi - \alpha_k k - \alpha_0$ provides an estimate of the unobserved productivity state ω which will be used as one of the arguments to control for the expectation in equation (3.6). So identification of α_k either requires controlling for the expectation without simultaneously controlling for k' or it has to come from including ω in the nonparametric function that controls for the expectation, so that identification of α_k comes from the index restrictions given by $\omega = \phi - \alpha_k k - \alpha_0$ for the resulting multiple index model.

The introduction of an unobserved knowledge state makes the task of identifying α_k more difficult as it calls for controlling for the second unobserved state variable. The question is whether R&D can serve as a proxy for the knowledge state, while preserving identification. Consider the optimal knowledge choice ψ' conditional on the optimal capital choice k',

$$\begin{aligned} \psi' &= f(\omega, k', \psi) \\ &= \arg\sup_{\psi'} -r(\psi', \omega, \psi) + \beta E(V(\omega', k', \psi')|\psi'). \end{aligned} \tag{3.7}$$

It is clear from this equation, that controlling for ψ' also requires a proxy for ψ. Unfortunately, we cannot use the result of Theorem 3.3 here since this would also control for ω which would leave α_k unidentified. Instead, an alternative way to proxy for ψ needs to be explored. The following theorem addresses this.

Theorem 3.5 *Assume that assumptions (A0) to (A4) hold, that the value function is twice continuously differentiable, and that one of the following additional* *conditions is satisfied everywhere*

$$\begin{aligned} \frac{\partial r}{\partial \psi'}\frac{\partial^2 r}{\partial \psi' \partial \psi} &< -\frac{\partial r}{\partial \psi}\frac{\partial^2}{\partial \psi'^2}[-r(\psi', \omega, \psi) + \beta E(V(\omega', k', \psi')|\psi')] \\ \text{or } \frac{\partial r}{\partial \psi'}\frac{\partial^2 r}{\partial \psi' \partial \psi} &> -\frac{\partial r}{\partial \psi}\frac{\partial^2}{\partial \psi'^2}[-r(\psi', \omega, \psi) + \beta E(V(\omega', k', \psi')|\psi')]. \end{aligned}$$

Then the unobserved knowledge choice ψ' can be written as a function of the productivity state ω, the capital choice k' and the observed R&D investments r,

$$\psi' = g(\omega, k', r).$$

Proof. *Consider the first order condition of the optimization problem in equation (3.7):*

$$-\frac{\partial}{\psi'}r(\psi', \omega, \psi) + \beta\frac{\partial}{\psi'}E(V(\omega', k', \psi')|\psi') = 0$$

Conditional on ω and k' this defines a schedule in (ψ', ψ) space along which the FOC holds. The slope of the schedule is

$$\frac{d\psi'}{d\psi} = \frac{\frac{\partial^2 r}{\partial \psi' \partial \psi}}{\frac{\partial^2}{\partial \psi'^2}[-r(\psi', \omega, \psi) + \beta E(V(\omega', k', \psi')|\psi')]} > 0.$$

The change in R&D investments implied by a movement along the schedule are given by the total differential of the R&D function holding ω fixed.

$$dr = \frac{\partial r}{\partial \psi'}d\psi' + \frac{\partial r}{\partial \psi}d\psi$$

For R&D to pin down which point along the schedule is consistent with the data, R&D must be either strictly monotonically increasing or strictly monotonically decreasing along this schedule. Substituting the slope of the schedule above into the differential yields the conditions of the theorem. ■

Note that the condition required for the theorem involves the second derivative of the future expected value of the firm. This is unfortunate, as the value function is, of course, endogenously determined by all the primitives of the model. Hence, we cannot check whether the conditions of the theorem hold (even post estimation) without specifying all the model primitives and solving the entire dynamic programme for the value function. Solving the dynamic programme, however, is the very complicated task which the semiparametric estimation approach tries to avoid in the first place. It is therefore impossible to judge whether the conditions for the theorem are likely to be satisfied in praxis. Absent any better alternatives to identify α_k, all we can do at this stage is to accept the conditions and assume that they are satisfied.[3]

Theorem 3.5 implies that, under the conditions of the theorem, the expectation in the second stage estimation equation can be rewritten as a nonparametric function of the productivity estimate $\phi - \alpha_k k$, the capital choice k', and observed R&D investments r,

$$E[\omega'|\psi', \chi' = 1] = h(\phi - \alpha_k k, k', r).$$

As in the previous chapter, this also controls for survival as controlling for ψ' and k' also controls for the exit threshold in the exit equation (3.2). This results in a fully nonparametric form for the stage two estimation equation

$$y^{*\prime} = \widetilde{h}(\phi - \alpha_k k, k', r) + \xi' + \eta', \tag{3.8}$$

where the identification of $\alpha_k k$ comes from the index restrictions on the nonparametric function $\widetilde{h}()$.

The original formulation of the second stage estimation in Olley & Pakes (1996) suggests that an alternative estimation equation for stage two which includes a separate estimate of the probability of survival P instead of k' to proxy for the expectation,

$$y^{*\prime} = \alpha_k k' + \widetilde{\widetilde{h}}(\phi - \alpha_k k, P, r) + \xi' + \eta'. \tag{3.9}$$

Recall from the previous chapter that the survival probability is strictly increasing in next period's knowledge ψ' and strictly decreasing in the exit threshold $\underline{\omega}(k', \psi')$. The exit threshold is nondecreasing in its arguments by the monotonicity of the

[3] Again, $\frac{\partial^2 r}{\partial \psi' \partial \psi} = 0$ would ensure that the conditions of the theorem are trivially satisfied. As argued above, including of R&D in the list of proxies for the expectation would then be unnecessary.

value function. Furthermore, an Euler equation argument for the optimal choice of ψ' conditional on ω, k', and ψ, coupled with the additional conditions for Theorem 3.5 establishes that the optimal knowledge choice ψ' conditional on ω, k', and r is strictly increasing in k'. Therefore, the survival probability is strictly increasing in k' conditional on ω and r so equations (3.8) and (3.9) should be asymptotically equivalent as in the previous chapter.

3.4 Production function estimates

3.4.1 Data and specification details

We estimate these specifications for the four 3-digit SIC industries "Pharmaceuticals (SIC 283)", "Computer Hardware (SIC 357)", "Telecommunications Equipment (SIC 366)", and "Computer Software (SIC 737)" using the dataset of the previous chapter. The data covers firms listed on North American stock markets over the years 1980 to 2001 and is constructed from the COMPUSTAT database.[4]

The nonparametric functions in the estimation equations (3.5), (3.8), and (3.9) are approximated by polynomial series expansions. For the stage one estimation of equation (3.5), the nonparametric function $\phi()$ is approximated by a fourth order polynomial in capital, investment, R&D, and time. Allowing time to enter this function accounts for changes in the economic environment over time leading to changes in the investment policy function. To control for the expected productivity conditional on survival, the second stage of the estimation procedure uses a third order polynomial expansion in the lagged productivity estimate, lagged R&D and either the current capital stock or lagged survival probability depending on whether equation (3.8) or (3.9) is estimated. The survival probability in equation (3.9) is estimated using a probit regression with a third order polynomials series expansion in capital, investment, R&D, and time. In both stages of the estimation we use a least squares approach to obtain estimates. The productivity estimate entering the nonparametric function of the second stage is, by equation (3.5), the stage one

[4]See Section 2.4 and Appendix A2.2 of the previous chapter for a description of the data, summary statistics, and details of the data construction.

estimate of the function $\phi()$ minus the capital contribution (which is a function of the capital coefficient).

We obtain standard errors by bootstrapping using 200 repetitions. We treat the firm as the unit of observations giving each firm equal probability to be selected. A particular bootstrap sample is considered complete when its number of firm year observations is equal to or just exceeds the number of observations in the original sample. We use the same bootstrap samples across specifications which enables us to generate distributions for the differences of parameter estimates across specifications.

3.4.2 Estimation results

Table 3.1 reports regression results for each of the four industries (panel a to d). For comparison, the first three columns in each panel repeat the estimation results for the model without unobserved knowledge of the previous chapter. Column "o" gives simple OLS estimates, column "k" corresponds to the fully nonlinear specification of the previous chapter using capital in stage two (equation (2.18)), and column "p" presents the (asymptotically equivalent) Olley-Pakes estimates using the survival probability (equation (2.17)). The last two columns present estimates for the specifications for the model with unobserved knowledge state in this chapter. Column "kr" corresponds to stage two equation (3.8) using the lagged productivity estimate, current capital, and lagged R&D as arguments of the nonparametric function, while column "pr" uses the lagged survival probability estimate instead of current capital corresponding to stage two equation (3.9).

In the last chapter, we discussed that the OLS estimate of the labour coefficient is expected to be biased upwards due to the positive correlation of labour with productivity. Indeed, the labour coefficient estimated for the model of the previous chapter is lower than the OLS estimate in each industry. This is still true in the model with unobserved knowledge of this chapter. However, it is interesting to note that in all industries except "Computers", the labour estimates for the model with unobserved knowledge capital lies between the OLS estimate and the estimate of the previous chapter. This implies that the OLS bias may be less serious than suggested by the results of the last chapter if the true model includes

Table 3.1: Production function estimates with unobserved knowledge

a. Pharmaceuticals (SIC 283)

Dep. Var: Value Added		OLS	No knowledge cap.		Knowledge capital	
			Stage 2 variables in addition to lagged omega:			
			Capital	L.Prob	Capital L.R&D	L.Prob L.R&D
		o	k	p	kr	pr
Labour		**0.748**	**0.687**	**0.687**	**0.701**	**0.701**
	SE	0.041	0.044	0.044	0.039	0.039
Capital		**0.388**	**0.304**	**0.378**	**0.299**	**0.277**
	SE	0.038	0.114	0.073	0.134	0.065
# obs		3932	3269	3269	3009	3009
# firms		461	461	461	434	434
RSS		456.7	316.0	316.8	284.9	285.1

b. Computer Hardware (SIC 357)

Dep. Var: Value Added		OLS	No knowledge cap.		Knowledge capital	
			Stage 2 variables in addition to lagged omega:			
			Capital	L.Prob	Capital L.R&D	L.Prob L.R&D
		o	k	p	kr	pr
Labour		**0.973**	**0.833**	**0.833**	**0.810**	**0.810**
	SE	0.037	0.043	0.043	0.044	0.044
Capital		**0.067**	**0.206**	**0.291**	**0.142**	**0.365**
	SE	0.035	0.076	0.097	0.057	0.139
# obs		2210	1866	1866	1816	1816
# firms		253	253	253	244	244
RSS		186.9	139.4	138.5	123.2	122.7

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.
All specifications other than OLS are labelled according to the variables included in stage 2 of the estimation procedure.
Specifications "k" and "p" for the model without knowledge capital use a polynomial expansion in investment, capital, and time of order 5 to proxy for productivity in stage1.
Specifications "kr" and "pr" for the model with knowledge capital use a polynomial expansion in investment, capital, R&D, and time of order 4 to proxy for productivity in stage1.
In all cases a 3rd order polynomial expansion is used in stage 2.
The OLS specification also includes a 5th order polynomial in time to control for time effects. Including time dummies yields very similar results.

c. Telecom Equipment (SIC 366)

Dep. Var:	OLS	No knowledge cap.		Knowledge capital	
		Stage 2 variables in addition to lagged omega:			
Value Added		Capital	L.Prob	Capital L.R&D	L.Prob L.R&D
	o	k	p	kr	pr
Labour	**0.826**	**0.688**	**0.688**	**0.701**	**0.701**
SE	0.038	0.038	0.038	0.036	0.036
Capital	**0.279**	**0.327**	**0.410**	**0.205**	**0.276**
SE	0.034	0.078	0.062	0.062	0.073
# obs	2443	2094	2094	1960	1960
# firms	259	259	259	253	253
RSS	203.0	158.0	156.6	134.6	134.1

d. Software (SIC 737)

Dep. Var:	OLS	No knowledge cap.		Knowledge capital	
		Stage 2 variables in addition to lagged omega:			
Value Added		Capital	L.Prob	Capital L.R&D	L.Prob L.R&D
	o	k	p	kr	pr
Labour	**0.775**	**0.684**	**0.684**	**0.727**	**0.727**
SE	0.031	0.037	0.037	0.035	0.035
Capital	**0.380**	**0.463**	**0.432**	**0.385**	**0.376**
SE	0.026	0.057	0.036	0.069	0.063
# obs	4526	3647	3647	3021	3021
# firms	693	693	693	622	622
RSS	387.5	239.1	238.8	180.3	181.1

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.
All specifications other than OLS are labelled according to the variables included in stage 2 of the estimation procedure.
Specifications "k" and "p" for the model without knowledge capital use a polynomial expansion in investment, capital, and time of order 5 to proxy for productivity in stage1.
Specifications "kr" and "pr" for the model with knowledge capital use a polynomial expansion in investment, capital, R&D, and time of order 4 to proxy for productivity in stage1.
In all cases a 3rd order polynomial expansion is used in stage 2.
The OLS specification also includes a 5th order polynomial in time to control for time effects. Including time dummies yields very similar results.

a knowledge state. Panel b of table 3.2 gives an indication of the significance of the differences. The numbers in the table report the fraction of cases in the 200 bootstrap samples plus one original sample in which the difference of the coefficients of the two specifications that are compared is positive. "OLS" refers to the OLS estimate, "No Kn." to the R&D model of the last chapter without knowledge capital and "Knowl." refers to the model with an unobserved knowledge state in this chapter. While the OLS labour coefficient is significantly higher than the estimates of the previous chapter in all industries, the difference for the estimates of this chapter is only significant for "Computer Hardware" and "Telecommunication Equipment". The comparison between the estimates of this and the last chapter shows that statistically, the labour coefficients are indistinguishable, with the exception of "Software", where the labour coefficient of this chapter is significantly higher than that of the last.

With respect to the estimates of the capital coefficient, the two (asymptotically equivalent) specifications of this chapter produce fairly similar estimates for "Pharmaceuticals" and for "Software", while specification "pr" seems to produce higher capital estimates than specification "kr" for "Computers" and "Telecom Equipment" (Table 3.1). Panel a of table 3.2 however reveals that this difference is statistically significant for "Computers" only.

Compared to the OLS estimates, the capital estimates of specification "kr" are statistically indistinguishable, even though the point estimates for "Pharmaceuticals" and "Telecom Equipment" of specification "kr" are lower than the OLS estimates. In contrast, the capital estimates of specification "pr" are significantly lower than the OLS estimate for "Pharmaceuticals" and significantly higher than OLS for "Computers".

Compared to the two specifications of the previous chapter, specification "kr" for the model with a knowledge state produces significantly lower estimates in the industries "Computers" and "Telecom Equipment". The estimates of specification "pr" are statistically indistinguishable from the estimates of the previous chapter.

In summary, it seems that while the estimated labour coefficients for the model with unobserved knowledge tend to lie between those of the model of the previous chapter and the OLS estimates, no obvious systematic pattern emerges for

Table 3.2: Differences in coefficient estimates with unobserved knowledge

a. Capltal coefficient

SIC	283	357	366	737
kr > pr	0.512	0.075	0.129	0.463
kr > o	0.134	0.816	0.174	0.473
kr > k	0.478	0.030	0.050	0.134
kr > p	0.294	0.000	0.025	0.189
pr > o	0.040	0.940	0.552	0.502
pr > k	0.542	0.871	0.398	0.100
pr > p	0.289	0.692	0.129	0.189
o > k	0.896	0.025	0.313	0.090
o > p	0.816	0.000	0.075	0.045
k > p	0.264	0.095	0.199	0.637

b. Labour coefficlent

SIC	283	357	366	737
OLS > No kn.	0.995	1.000	1.000	1.000
OLS > Knowl	0.851	1.000	1.000	0.896
No kn.>Knowl.	0.294	0.796	0.244	0.060

Each row of the table gives a comparison of the capital coefficient for two specifications. The number in each cell reports the percentage of cases across the 201 samples (one original sample + 200 bootstrapped samples) in which the capital coefficient in the first specification exceeds that of the second specification.
E.g. the first entry for row "kr>pr" implies that in 51.2% of the samples for Pharmaceuticals, the estimate of the capital coefficient in pecification "kr" exceeds that in specification "pr". In panel b, "No kn." stands for the model without knowledge capital and "Knowl." for the model with knowledge capital.

the estimates of the capital coefficients and the differences are for the most part statistically insignificant.

3.4.3 Specification tests

We now turn to test whether the model of this chapter is appropriate for the industries under study. The logic for the test is the same as in Olley & Pakes (1996) and in the previous chapter. If the investment and R&D are sufficient to proxy for the unobserved state variables and the model is specified correctly, then including lagged values of the observed labour and capital inputs linearly in the second stage of the estimation should not affect the estimation results and the coefficient estimates for these variables should be zero. Table 3.3 reports the results from these regressions. The table shows that lagged labour has a significantly negative coefficient in all industries except "Pharmaceuticals". Therefore, we conclude that the model is rejected for the industries "Computers", "Telecommunications Equipment" and "Software".

Allowing for an unobserved knowledge state by including R&D in the list of proxies does hence not improve on the model rejections of the previous chapter. In fact, for "Telecommunications Equipment" specification "p" of the previous chapter was accepted, while none of the two specifications of this chapter is. This is surprising, as the model of this chapter is more general than that of the previous chapter.

On the other hand, the results of the previous chapter show that when R&D is included in the second stage as a proxy (but not in stage one), the resulting capital estimates are unreasonable low. This may be an indication, that R&D is in fact proxying for unobserved knowledge capital as proposed in this chapter and that it should enter stage 1 of the estimation.

It is also interesting to note, that the model rejections are due to a significant coefficient on lagged labour. This suggest that maybe labour is not a completely variable factor (for example due to adjustment costs of the labour force) but should be treated as a state variable in the model.[5]

[5] Conceptionally, including a labour state in the model is straightforward as the monotonicity results would not be substantially affected. Both coefficients would then have to be estimated

Table 3.3: Specification tests for the model with unobserved knowledge

a. Pharmaceuticals (SIC 283)

Dep. Var:
Value Added

	pr	pr	kr
Labour	**0.701**	**0.701**	**0.701**
SE	0.039	0.039	0.039
Capital	**0.287**	**0.288**	**0.300**
SE	0.070	0.071	0.133
L.Labour		**-0.014**	**-0.025**
SE		0.018	0.023
fraction>0		0.418	0.274
L.Capital	**-0.019**		
SE	0.023		
fraction>0	0.388		

Numbers below coefficients are standard errors from 200 bootstrap sample sand the percentage of cases with positive coefficients

b. Computers (SIC 357)

Dep. Var:
Value Added

	pr	pr	kr
Labour	**0.810**	**0.810**	**0.810**
SE	0.044	0.044	0.044
Capital	**0.365**	**0.413**	**0.163**
SE	876.946	0.149	0.059
L.Labour		**-0.104**	**-0.113**
SE		0.033	0.022
fraction>0		0.000	0.000
L.Capital	**-0.047**		
SE	433654.5		
fraction>0	0.065		

Numbers below coefficients are standard errors from 200 bootstrap sample sand the percentage of cases with positive coefficients

c. Telecom Eqmt (SIC 366)

Dep. Var:
Value Added

	pr	pr	kr
Labour	**0.701**	**0.701**	**0.701**
SE	0.036	0.036	0.036
Capital	**0.298**	**0.341**	**0.199**
SE	0.082	0.087	0.063
L.Labour		**-0.042**	**-0.070**
SE		0.018	0.019
fraction>0		0.005	0.000
L.Capital	**-0.013**		
SE	0.020		
fraction>0	0.234		

Numbers below coefficients are standard errors from 200 bootstrap sample sand the percentage of cases with positive coefficients

d. Software (SIC 737)

Dep. Var:
Value Added

	pr	pr	kr
Labour	**0.727**	**0.727**	**0.727**
SE	0.035	0.035	0.035
Capital	**0.387**	**0.487**	**0.384**
SE	0.060	0.064	0.070
L.Labour		**-0.094**	**-0.134**
SE		0.024	0.024
fraction>0		0.000	0.000
L.Capital	**-0.022**		
SE	0.015		
fraction>0	0.109		

Numbers below coefficients are standard errors from 200 bootstrap sample sand the percentage of cases with positive coefficients off the index restrictions of the nonparametric functionl in stage two which could be done in a GMM framework similar to the one in Levinsohn & Petrin (2003) for the no selection, no R&D, no knowledge capital case.

3.5 Knowledge, R&D, and the future productivity distribution

The productivity estimates of the previous section can now be used to analyse the effect of current productivity, R&D, and knowledge capital on the distribution of next period's productivity. Knowledge capital is, of course, an unobserved state variable so it has to be estimated first. To do so, note that the distribution of future productivity is, according to the assumptions of our model, first order stochastically increasing in the unobserved future knowledge state. So any statistic of this distribution that increases as the distribution increases in the FOSD sense, e.g. the mean of the distribution, is a sufficient statistic for next period's knowledge, as the knowledge state must be a strictly monotone transformation of this statistic.

Moreover, the invertibility result of Theorem 3.3 together with the policy function (3.4) imply that next period's knowledge state is a function of the current capital stock, investment and R&D. The difficulty in estimating the knowledge stock on the basis of the productivity distribution lies in the selection problem, as the productivity distribution conditional on survival not only depends on the knowledge state, but also on the exit threshold. For reasons of simplicity, we will ignore this issue in this section.

Therefore, if we are willing to abstract from selection issues, we can obtain an estimate of the current knowledge state by nonparametrically regressing current productivity on lagged capital, lagged investment, and lagged R&D.[6] We use OLS with a third order polynomial expansion for this regression. The prediction from this regression is a sufficient statistic for the unobserved knowledge state which we use in subsequent regressions.

Table 3.4 presents the results of OLS and quantile regressions of future productivity on current productivity ω, knowledge ψ, and log R&D for each industry and using the productivity estimates from each of the specifications "kr" and "pr" of the previous section. The regressions are clearly simplistic in that they specify a simple linear functional form. Standard errors are based on the same 200 boot-

[6] As in the previous chapter, we remove time effects from the productivity estimates by using deviations from the sales weighted industry-year means.

strap samples as in the previous section and both the productivity estimates and the knowledge capital are re-estimated for each bootstrap sample thereby taking into account the variance in these estimates across samples.

Let us first examine the first order stochastic dominance assumption of the theoretical model. For "Pharmaceuticals" and "Telecom Equipment", the stochastic dominance assumptions seem to be satisfied, as the coefficients for productivity, knowledge, and R&D on all the quantiles are positive with some significantly so. On the other hand, the assumption seems to be rejected for "Computers" and "Software". This is the same result as for the model without knowledge capital.

With respect to the effect of R&D on the distribution, the results of the model without knowledge capital from the previous chapter also seem relatively robust. For "Pharmaceuticals", a crude estimate of the short run elasticity of productivity with respect to R&D is around .016 (respectively .17 for specification "pr") whereas it is about between .028 and .036 for "Telecom Equipment". In the model without knowledge capital, these estimates are around .02 for "Pharmaceuticals" and between .007 and .04 for "Telecom Equipment" (Tables 2.9 and 2.11). The estimates with knowledge are significant as Table 3.5 reveals. For "Computers" and "Software" the point estimates are negative but insignificant, confirming the rejection of the model for these industries.

Of key interest for the model of this chapter is, however, the quantitative importance of the effect of the knowledge state ψ on the future productivity distribution relative to the current productivity state ω. For "Pharmaceuticals" the coefficient on knowledge in the OLS regressions is of similar quantitative importance than the current productivity realisation (around .4 for knowledge and around .5 for productivity). Moreover, the effect of knowledge on the mean and on all the quantiles is highly significant (Table 3.6). This suggests that knowledge capital is of quantitative importance in this industry. As knowledge capital allows, to some extent, for lagged effects of R&D on productivity, this is in line with the prior, that there are typically long lags in the Pharmaceutical industry between the innovation of a new drug and bringing it to the market.

On the other hand, for "Computers", "Telecom Equipment", and "Software" the effect of knowledge capital seems to be much weaker. Point estimates are very

Table 3.4: Quantile regressions on the conditional productivity distribution for the model with unobserved knowledge

Dependent variable: Omega(t+1)

a. Pharmaceuticals (SIC 283)

	Specification "kr"				Specification "pr"			
F.Omega	**OLS**	**Quantile**			**OLS**	**Quantile**		
		0.25	**0.5**	**0.75**		**0.25**	**0.5**	**0.75**
Omega	**0.486**	**0.674**	**0.733**	**0.740**	**0.501**	**0.707**	**0.760**	**0.744**
SE	0.091	0.110	0.099	0.088	0.062	0.079	0.071	0.070
Knowledge	**0.402**	**0.258**	**0.164**	**0.101**	**0.392**	**0.235**	**0.141**	**0.099**
SE	0.087	0.090	0.092	0.100	0.065	0.072	0.070	0.077
Log_R&D	**0.016**	**0.018**	**0.017**	**0.013**	**0.017**	**0.018**	**0.018**	**0.016**
SE	0.009	0.007	0.009	0.014	0.006	0.006	0.006	0.010
const	**-0.137**	**-0.176**	**-0.143**	**-0.100**	**-0.148**	**-0.174**	**-0.154**	**-0.127**
SE	0.078	0.064	0.075	0.123	0.053	0.055	0.055	0.085
pseudo R^2	0.771	0.612	0.660	0.661	0.809	0.641	0.690	0.696
# obs	2494	2494			2494	2494		

b. Computer Hardware (SIC 357)

	Specification "kr"				Specification "pr"			
F.Omega	**OLS**	**Quantile**			**OLS**	**Quantile**		
		0.25	**0.5**	**0.75**		**0.25**	**0.5**	**0.75**
Omega	**0.879**	**0.884**	**0.917**	**0.911**	**0.759**	**0.763**	**0.795**	**0.786**
SE	0.061	0.063	0.061	0.072	0.089	0.099	0.090	0.091
Knowledge	**-0.006**	**-0.036**	**-0.009**	**0.020**	**0.111**	**0.056**	**0.082**	**0.166**
SE	0.068	0.084	0.069	0.079	0.113	0.118	0.103	0.106
Log_R&D	**-0.001**	**-0.001**	**-0.002**	**-0.004**	**-0.020**	**-0.024**	**-0.022**	**-0.012**
SE	0.005	0.008	0.005	0.006	0.009	0.013	0.010	0.008
const	**-0.026**	**-0.079**	**-0.012**	**0.051**	**0.146**	**0.127**	**0.158**	**0.128**
SE	0.036	0.062	0.041	0.046	0.077	0.107	0.083	0.066
pseudo R^2	0.703	0.483	0.498	0.472	0.808	0.604	0.600	0.563
# obs	1528	1528			1528	1528		

Numbers below coefficients are standard errors of the coefficients across 200 bootstrapped samples from the production function estimation.

Dependent variable: Omega(t+1)

c. Telecom Equipment (SIC 366)

	Specification "kr"				Specification "pr"			
F.Omega	**OLS**	**Quantile**			**OLS**	**Quantile**		
		0.25	**0.5**	**0.75**		**0.25**	**0.5**	**0.75**
Omega	**0.714**	**0.781**	**0.776**	**0.674**	**0.681**	**0.752**	**0.729**	**0.649**
SE	0.087	0.090	0.078	0.091	0.086	0.088	0.078	0.085
Knowledge	**0.103**	**0.035**	**0.101**	**0.145**	**0.114**	**-0.029**	**0.070**	**0.232**
SE	0.096	0.094	0.077	0.112	0.106	0.091	0.082	0.112
Log_R&D	**0.036**	**0.037**	**0.027**	**0.037**	**0.028**	**0.037**	**0.028**	**0.020**
SE	0.012	0.015	0.010	0.016	0.010	0.014	0.011	0.015
const	**-0.308**	**-0.370**	**-0.226**	**-0.259**	**-0.235**	**-0.368**	**-0.239**	**-0.108**
SE	0.103	0.124	0.088	0.136	0.090	0.122	0.096	0.121
pseudo R^2	0.820	0.577	0.601	0.615	0.740	0.496	0.529	0.539
# obs	1670	1670			1670	1670		

d. Software (SIC 737)

	Specification "kr"				Specification "pr"			
F.Omega	**OLS**	**Quantile**			**OLS**	**Quantile**		
		0.25	**0.5**	**0.75**		**0.25**	**0.5**	**0.75**
Omega	**0.776**	**0.788**	**0.862**	**0.869**	**0.781**	**0.796**	**0.868**	**0.873**
SE	0.083	0.097	0.090	0.082	0.081	0.096	0.088	0.080
Knowledge	**0.120**	**0.073**	**0.035**	**0.071**	**0.150**	**0.108**	**0.040**	**0.075**
SE	0.105	0.131	0.096	0.085	0.081	0.139	0.100	0.085
Log_R&D	**-0.003**	**0.004**	**0.000**	**-0.005**	**-0.004**	**0.004**	**0.002**	**-0.005**
SE	0.012	0.012	0.009	0.009	0.008	0.011	0.008	0.008
const	**-0.002**	**-0.091**	**-0.027**	**0.048**	**0.003**	**-0.088**	**-0.035**	**0.048**
SE	0.093	0.090	0.069	0.067	0.057	0.083	0.057	0.058
pseudo R^2	0.666	0.451	0.524	0.528	0.674	0.462	0.530	0.532
# obs	2334	2334			2334	2334		

Numbers below coefficients are standard errors of the coefficients across
200 bootstrapped samples from the production function estimation.

Table 3.5: Significance of the effect of R&D on the productivity distribution

a. Specification "kr" - R&D

R&D	SIC	OLS	Quantile 0.25	0.5	0.75
Pharmaceut.	283	0.988	0.995	0.978	0.913
Computers	357	0.634	0.709	0.644	0.595
Telecom Eqmt	366	0.995	0.983	0.993	0.968
Software	737	0.607	0.756	0.706	0.582

b. Specification "pr" - R&D

R&D	SIC	OLS	Quantile 0.25	0.5	0.75
Pharmaceut.	283	1.000	1.000	1.000	0.968
Computers	357	0.540	0.577	0.540	0.532
Telecom Eqmt	366	0.985	0.985	0.990	0.900
Software	737	0.602	0.734	0.704	0.580

The number in each cell reports the percentage of cases across the 201 samples (one original sample + 200 bootstrapped samples) with a positive coefficient on R&D

Table 3.6: Significance of the effect of Knowledge on the productivity distribution

a. Specification "kr" - Knowledge capital

Psi	SIC	OLS	Quantile 0.25	0.5	0.75
Pharmaceut.	283	1.000	0.993	0.973	0.955
Computers	357	0.900	0.856	0.886	0.878
Telecom Eqmt	366	0.975	0.846	0.943	0.985
Software	737	0.960	0.930	0.903	0.925

b. Specification "pr" - Knowledge capital

Psi	SIC	OLS	Quantile 0.25	0.5	0.75
Pharmaceut.	283	1.000	1.000	1.000	0.993
Computers	357	0.948	0.861	0.920	0.955
Telecom Eqmt	366	0.978	0.803	0.948	0.990
Software	737	0.963	0.930	0.910	0.928

The number in each cell reports the percentage of cases across the 201 samples (one original sample + 200 bootstrapped samples) with a positive coefficient on Psi close to zero in "Computers" and around .15 in "Telecom Equipment" and "Software" (compared to coefficients on productivity that are well above .7). While these coefficients are still significant, these results suggest that lagged effects of R&D on productivity through knowledge are quantitatively relatively less important than in "Pharmaceuticals" and that the simple view of the purely stochastic accumulation of R&D into productivity of the previous chapter may not be a bad assumption for these industries – at least for "Telecom Equipment" for which we accepted the model of the previous chapter in the specification tests.

3.6 Conclusion

This chapter introduces unobserved knowledge capital into the empirical framework of the previous chapter. This is an important extension, as it allows us to reconcile two extreme views on the accumulation of R&D in the literature. Moreover, it also allows for the possibility of lagged effects of R&D on productivity through unobserved knowledge. Methodologically, we show that the invertibility approach by Olley & Pakes (1996) can still be applied to this model with two unobserved states and that, under certain assumptions, observed R&D investments can serve as an additional proxy for the second state variable in the estimation.

For the industries in our study, the production function estimates seem to suggest that the OLS biases in the labour coefficient are somewhat less severe as suggested by the Olley & Pakes (1996) approach for the model without unobserved knowledge. However, the change in the parameter estimates due to the introduction of unobserved knowledge is moderate and for the most part insignificant.

The results on the short run elasticity of productivity with respect to R&D are very similar to that of the model without knowledge. Therefore, the empirical results of the previous chapter are for the most part robust to the introduction of unobserved knowledge.

Interestingly, the effect of knowledge capital on future productivity is most pronounced for "Pharmaceuticals". As knowledge capital captures the effect of lags in the R&D process, this is very plausible as there are typically long time-lags between the innovation of a new drug and the time when it becomes available on the market (i.e. when they result in a payoff).

Unfortunately, the model is rejected for the three industries "Computer Hardware", "Telecommunications Equipment" and "Software". For these industries, using investment and R&D as joint proxies does not seem to be sufficient to capture unobserved productivity and knowledge differences across firms.

Conclusion

The most striking empirical observation of this thesis is the "negative tail" of small loss-making firms in the cross sectional profits-size distribution (Figure 1-1) and the fact that this tail is significantly stronger in high R&D industries than in low R&D industries. Since the cross sectional distribution is, of course, an endogenous outcome of the underlying dynamics, this thesis analyses different complementary dynamic models of firm profitability, growth, and exit.

Chapter 1 first documents the empirical regularities in the cross sectional profits-size distribution and the intra-distribution dynamics, and the variation across industries. The fact that the negative tail of small-loss making firms is systematically stronger in high R&D industries than in low R&D industries suggests that there are two strong underlying mechanisms at work that are common across industries: a real option effect explaining the presence of small, highly unprofitable firms in the distribution; and a diversification effect which explains why very few large firms are unprofitable.

The chapter the proposes a simple stochastic reduced form model for firm dynamics that combines these two effects. In this model, firms consist of independent businesses of the same size that arrive randomly over time and whose profit flows evolve stochastically. The only decisions by the firms are decisions to optimally shut down businesses taking into account the option value of a loss-making business.

This model can generate the qualitative features of the profits-size distribution and the variation across high and low R&D industries by varying the variance in the underlying stochastic process. However, the scale of the size dimension in the model which is simply the count of active businesses of a firm does not agree with that in the empirical analysis where size is measured on a logarithmic scale. In this respect, the model is clearly not completely satisfactory. One potential research trajectory to address this would be to allow for individual businesses to be of different size and/or have different growth rates.

The model also fails to reproduce the most interesting stylized facts about intra-distribution dynamics. In particular, it does not generate the empirical observation that small unprofitable firms in high R&D industries are more likely to remain unprofitable than firms in low R&D industries. Given the simplicity of the model and its reduced form nature, this is perhaps not surprising. Industry dynamics are certainly more complex than the simple model which abstracts from many firm decisions such as investments in physical capital and R&D. Modelling these firm decisions and estimating underlying stochastic processes driving them is a challenging task which the second part of the thesis addresses.

higher sunk cost?

However, we believe that despite its failures, the fact that the simple model generates the qualitative differences in the profits-size distributions between high and low R&D industries by varying the riskiness of the economic environment is a strong indication the empirical features are due to a combination of a real option and a diversification effect.

The second part of the thesis considers a much richer structural model for firm dynamics. This model abstracts from the diversification effect, but explicitly models firms investment decisions in physical capital and in R&D. In this part, the focus of the thesis lies on developing techniques for the estimation of the central stochastic primitive in the recent theoretical literature of firm dynamics that drives the dynamics of firm profitability, growth, and exit: The distribution of future productivity conditional on the current state and on R&D investments.

Since productivity is typically an unobserved state variable in these models, estimation of the conditional productivity distribution requires techniques that recover this unobserved state variable from observed variables. To do so, we follow the invertibility approach originally proposed by Olley & Pakes (1996). This approach is based on the idea that if the investment policy function of the model is strictly monotonic in productivity, there is a unique mapping from the observed states and the observed investment choices into productivity. This allows to control for unobserved productivity semiparametrically in the estimation of production functions. This estimation also produces estimates of the unobserved productivity state which can the be used to analyse the conditional productivity distribution.

Chapter 2 introduces R&D into the model of Olley & Pakes (1996) and proves that the investment policy function for physical capital satisfies the monotonicity property. While the presence of R&D requires the researcher to control for expected productivity conditional on past R&D, the original Olley-Pakes estimation equation indirectly takes account of this by including the survival probability in the estimation. The chapter also presents an alternative estimation equation which directly addresses this issue and which transforms the second stage estimation into a fully nonlinear multiple index model. The productivity estimates from the production function estimation are then used to test the model, and to estimate the conditional productivity distribution and the short run returns to R&D. ?

Chapter 3 takes one step further in that it introduces a second unobserved state variable in the model, an unobserved knowledge state. This reconciles the Griliches-type accumulation of R&D into knowledge capital and the stochastic accumulation of R&D into productivity a la Ericson-Pakes. It also controls to some extent for the effect of lags in the R&D process. Methodologically, the chapter shows that the invertibility approach can still be applied in the presence of two unobserved state variables, provided R&D is included in the list of productivity proxies. distribution of knowledge eq?

Empirically the models are rejected for the Computer Hardware and Software industries and accepted for Pharmaceuticals. For Telecom Equipment the model of Chapter 2 is accepted but that of Chapter 3 is rejected. The chapters find short run elasticities of productivity with respect to R&D in the order of 2-3%. Furthermore, results indicate that allowing for lagged effects of R&D through knowledge is potentially important for Pharmaceuticals which confirms the prior of long lags in the R&D process for this industry.

The advantage of the semiparametric approach in the second part is that it is relatively easy to implement and computationally not very intensive. It also does not require functional form assumption for the productivity distribution or for other primitives of the model. Instead, all it requires are assumption on certain properties of primitives coupled with a functional form assumption for the production function. The semiparametric analysis and techniques of this part could be developed further along several dimensions.

First, to address the rejection of the model for the Computer and Software industries, labour could be included as a state variable in the model. If there are adjustment costs to labour such as hiring and firing costs, labour would indeed cease to be a completely variable factor of production and become a state. Conceptionally, including an observed labour state is straightforward and the monotonicity results would not be substantially affected. Empirically however, one would have to identify both the capital and the labour coefficient from the index restrictions in the second stage of the estimation while simultaneously controlling for R&D and selection. This places a much higher burden on the data.

Second, the fully nonlinear version of the stage two estimation equation lends itself to an alternative interpretation. It is essentially a nonparametric regression of future output on the current state coupled with index restriction to identify the coefficients of interest. One could also regress the future realisation of any other variable (such as labour, investment, capital, or R&D) on the state without introducing any additional assumptions in the model. This would lead to additional estimation equations in stage two which would improve the efficiency of the estimation procedure.

Third, the estimates generated by the semiparametric approach could be used to answer a range of additional questions. For example the knowledge capital estimates could serve to analyse the accumulation of knowledge through past knowledge, productivity, and R&D. This would shed light on an additional primitive of the dynamic model, the R&D function. One could also examine the depreciation of knowledge over time.

Finally, the most interesting question is whether the models of Chapters 2 and 3 can generate the empirical regularities documented in Chapter 1 – particularly the negative tail in the profits-size distribution – and, if so, whether they do better than the simple model of Chapter 1. The analysis in the second part of the thesis provides only a first step in this direction.

This is because of the semiparametric nature of the approach. While simple and computationally not very intensive, this approach falls short of estimating all the primitives of the dynamic model. To generate dynamics, however, a solution to the fully specified dynamic model is required. First, all its primitives would have to be specified and estimated. Then, one could solve the model numerically and use the generated policy functions to simulate dynamics. Finally, these simulations could be used to analyse the implied profits-size distribution.

Such an estimation could, for example, be implemented using a nested fixed point approach which solves the fully specified model numerically for each value of the parameter vector and maximises the fit of the model over the entire parameter vector. It is important to note that the invertibility results of Chapters 2 and 3 could still play a central role in such an approach as they provide a basis to control for unobserved state variables regardless of the estimation approach. However, a nested fixed point estimation is computationally very demanding and is beyond the scope of this thesis.

Appendix

Appendix to Chapter 1

A1.1 Data

The dataset is constructed from Standard and Poor's COMPUSTAT database on active and inactive ("research") companies over the period 1990 - 2002. The variables employed in this study are "Operating Income Before Depreciation" (CS data item A13 – OIBDP) for profits, the book value of "Total Assets" (CS data item A6 – AT), the firms' sales revenue (CS data item A12 – SALE), the number of employees (CS data item A29 – EMP), "Research and Development Expenditure" (CS data item A49 – XRD), and the firms "Advertising Expense" (CS data item A45 – XAD). Firms are allocated to industries according to their reported "Primary SIC code" (CS data item SIC). For firms that exit the dataset (inactive companies), the reason of deletion from the database is recorded in CS data item DLRSN. Where appropriate, variables are adjusted for inflation using the US GDP deflator.

The sample of 42 4-digit SIC industries is selected as follows. For the analysis of industry distributions, a certain number firm-year observations for each industry is required. First, all firm-year observations with missing profits, sales or asset data are deleted. Then the number of firms and firm-year observations in each industry is computed and we retain all industries with 50 or more firms and 300 or more firm years. Finally, industries with names containing "miscellaneous" (MISC) or "not elsewhere classified" (NEC) were dropped. This results in the sample of 42 industries.

For each industry, the R&D and advertising intensity is constructed as the (inflation adjusted) sum of R&D (resp. advertising) expenditures across firms and years divided by the (inflation adjusted) sum of sales across firms and years. Missing R&D or advertising data are treated as zeros.[7]

A1.2 The simple Brownian motion model

With the state variable ω being equal to the profit flow, the Bellman equation for the value function becomes:

$$v(\omega) = \max\{0, \omega dt + \frac{1}{1+\rho dt} E[v(\omega + d\omega)|\omega]\}.$$

When ω follows a Brownian motion with drift rate μ and variance parameter σ,

$$d\omega = \mu dt + \sigma dz$$

with dz being the increment of a standard Wiener process, the Bellman equation can be rewritten in the continuation region as

$$\begin{aligned} v(\omega) &= \omega dt + \frac{1}{1+\rho dt} E[v(\omega + d\omega)|\omega] \\ &= \omega dt + \frac{1}{1+\rho dt} E[v(\omega) + v'(\omega)d\omega + \frac{1}{2}v''(\omega)(d\omega)^2 + ...)|\omega] \\ &= \omega dt + (1-\rho dt)[v(\omega) + \mu v'(\omega)dt + \frac{1}{2}\sigma^2 v''(\omega)dt + ...)]. \end{aligned}$$

The second line uses a Taylor expansion and the third uses Ito's lemma. Rearranging, dividing by dt and letting $dt \to 0$ yields the second order differential equation:

$$-\rho v(\omega) + \mu v'(\omega) + \frac{1}{2}\sigma^2 v''(\omega) = -v.$$

To solve this equation, additional conditions are required. These are the the value matching and smooth pasting conditions which hold at the optimal stopping point

[7] In earlier versions, we only used data with nonmissing R&D data to construct R&D intensity. The qualitative results remained the same.

ω^*, and the condition that the marginal value of an increase in the state is bounded:

$$v(\omega^*) = 0; \quad v'(\omega^*) = 0; \quad 0 < \lim_{\omega \to \infty} v'(\omega) < \infty.$$

A particular solution of the differential equation takes the form $v_p(\omega) = c_1\omega + c_2$. Substitution into the differential equation yields $-\rho(c_1\omega + c_2) + \mu c_1 = -\omega$, and therefore $c_1 = \frac{1}{\rho}$ and $c_2 = \frac{\mu}{\rho^2}$. So the particular solution is

$$v_p(\omega) = \frac{1}{\rho}\omega + \frac{\mu}{\rho^2}.$$

The complementary function is a solution of the homogenous differential equation

$$-\rho v(\omega) + \mu v'(\omega) + \frac{1}{2}\sigma^2 v''(\omega) = 0.$$

An obvious candidate solution is $v_c(\omega) = Ae^{r\omega}$ where substitution into the homogenous differential equation shows that r has to solve

$$(-\rho + \mu r + \frac{1}{2}\sigma^2 r^2) = 0.$$

The roots for this equation are

$$r_1 = \frac{-\mu + \sqrt{\mu^2 + 2\sigma^2\rho}}{\sigma^2} > 0, r_2 = \frac{-\mu - \sqrt{\mu^2 + 2\sigma^2\rho}}{\sigma^2} < 0.$$

Therefore the solution for the complementary function takes the general form

$$v_c(\omega) = A_1 e^{r_1\omega} + A_2 e^{r_2\omega}.$$

The general solution is the sum of the complementary solution and the particular integral:

$$v(\omega) = v_c(\omega) + v_p(\omega) = A_1 e^{r_1\omega} + A_2 e^{r_2\omega} + \frac{1}{\rho}\omega + \frac{\mu}{\rho^2}.$$

The constants A_1 and A_2 and the value of ω^* are determined by substitution into the smooth pasting, value matching and the additional condition above

$$\begin{aligned} A_1 &= 0 \\ A_2 &= -\frac{1}{r_2\rho}e^{-r_2a} \\ \omega^* &= \frac{1}{r_2} - \frac{\mu}{\rho}. \end{aligned}$$

Therefore, the general solution is

$$\begin{aligned} v(\omega) &= -\frac{1}{r\rho}e^{r(\omega-\omega^*)} + \frac{1}{\rho}\omega + \frac{\mu}{\rho^2} \\ \omega^* &= \frac{1}{r} - \frac{\mu}{\rho} \\ \text{with } r &= -\frac{\mu + \sqrt{\mu^2 + 2\sigma^2\rho}}{\sigma^2} < 0. \end{aligned}$$

A1.3 The geometric Brownian motion model

Let the profit function for the model be

$$\pi(\omega_i) = 1 + \omega_i,$$

where $\omega_i \in (-\infty; 0]$ and where $-\omega_i$ can be interpreted as a measure of the operating costs of the business. Define

$$c \equiv -\omega_i$$

and let c follow a geometric Brownian motion with drift rate μ, variance σ^2, i.e.

$$\frac{dc}{c} = \mu dt + \sigma dz$$

where dz denotes the increments of a standard Wiener processes ($E(dz) = 0$, $E(dz^2) = dt$).

In the continuation region, the Bellman equation (1.2) now becomes

$$\begin{aligned} u(c) &= (1-c)dt + \frac{1}{1+\rho dt} E[u(c+dc)|c] \\ &= (1-c)dt + (1-\rho dt)E[u + u'dc + \frac{1}{2}u''dc^2 + ...|c] \\ &= (1-c)dt + (1-\rho dt)(u + u'\mu c dt + \frac{1}{2}u''\sigma^2 c^2 dt + ...) \end{aligned}$$

where $u(c) \equiv v(-c)$. The second line uses a Taylor expansion and the last line uses Ito's lemma. Rearranging, dividing by dt and letting $dt \rightarrow 0$ yields the second order partial differential equation corresponding to (1.6):

$$\rho u - \mu c u' - \frac{1}{2}\sigma^2 c^2 u'' = 1 - c$$

The value matching and smooth pasting conditions at the optimal stopping point c^* are:

$$u(c^*) = 0, u'(0) = 0$$

When c reaches zero, it remains there and the value of the business becomes just the annuity value of receiving a unit profit flow each period, hence:

$$v(0) = \frac{1}{\rho}$$

A particular solution to the partial differential equation takes the form $u_p(c) = \gamma_1 c + \gamma_2$. Substitution into the differential equation yields $\rho(\gamma_1 c + \gamma_2) - \mu c \gamma_1 = 1 - c$ and therefore $\gamma_1 = \frac{1}{\mu - \rho}$ and $\gamma_2 = \frac{1}{\rho}$. So the particular solution is:

$$u_p(c) = \frac{1}{\mu - \rho}c + \frac{1}{\rho}.$$

The complementary function is a solution of the homogenous differential equation

$$\rho u - \mu c u' - \frac{1}{2}\sigma^2 c^2 u'' = 0$$

An obvious candidate solution is $u_c(x) = Ac^r$, so r has to satisfy

$$\rho - (\mu - \frac{1}{2}\sigma^2)r - \frac{1}{2}\sigma^2 r^2 = 0.$$

The roots for this equation are:

$$\begin{aligned}
r_1 &= \frac{1}{2\sigma^2}\left(\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}\right) > 0, \\
r_2 &= \frac{1}{2\sigma^2}\left(\sigma^2 - 2\mu - \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}\right) < 0
\end{aligned}$$

Therefore the solution for the complementary function takes the general form

$$u_c(c) = A_1 c^{r_1} + A_2 c^{r_2}.$$

The general solution for the differential equation is the sum of the complementary solution and the particular integral:

$$u(c) = A_1 c^{r_1} + A_2 c^{r_2} + \frac{1}{\mu - \rho}c + \frac{1}{\rho},$$

where the constants A_1 and A_2 and the location of the optimal stopping point c^* still have to be determined. As c goes to zero, the value of the business has to go to $\frac{1}{\rho}$. As r_2 is negative, this implies that $A_2 = 0$. Knowing this, A_1 and c^* can be determined through substitution into the value matching and smooth pasting conditions, which yields the general solution

$$\begin{aligned}
u(c) &= \frac{a^{-r_1}}{\rho\,(r_1 - 1)}c^{r_1} + \frac{1}{\mu - \rho}c + \frac{1}{\rho} \\
\text{with } c^* &= \frac{\rho - \mu}{\rho}\frac{r_1}{r_1 - 1} \\
\text{and } r_1 &= \frac{1}{2\sigma^2}\left(\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\sigma^2\rho}\right) > 0
\end{aligned}$$

This is the solution in the text with $\omega = -c$, $v(\omega) = u(c)$, and $\omega^* = -c^*$.

Appendix to Chapter 2

A2.1 Proofs

Lemma 2.1 *The value function $V(\omega, k)$ of the model in equation (2.1) is*

1. *bounded under (A0),(A1),*
2. *nondecreasing in productivity ω and k under (A0),(A2),*
3. *unique under (A0),(A1), and*
4. *supermodular under (A0) to (A3).*

Proof. 1. Boundedness: *The boundedness of the profit function and the investment cost and the nonnegativity of R&D expenditure ensure that the one period returns are bounded from above. The fact that $\beta < 1$ then implies that the expected net present value of future one period returns is bounded above. Furthermore, Φ puts a lower bound on the value function so that $V(\omega, k)$ is bounded.*

2. Nondecreasing: *Since $\pi(\omega, k)$ is increasing and $r(\psi', \omega)$ decreasing in ω (and the expectation term independent of ω), the value function conditional on continuation*

$$V_S(\omega, k) = \sup_{k', \psi'} \left[\pi(\omega, k) - c(k', k) - r(\psi', \omega) + \beta \int V(\omega', k') dF(\omega'|\psi') \right]$$

is nondecreasing in ω. Hence $V(\omega, k) = \max[\Phi, V_S(\omega, k)]$ is nondecreasing in ω. The argument for monotonicity in k depends on the assumption on $c(k', k)$. If $c(k', k)$ is decreasing in k, the same argument as for the monotonicity in ω applies as in this case $\pi(\omega, k) - c(k', k)$ is increasing in k for every k' and hence $V(\omega, k)$ is increasing in k.

Under the alternative assumption that $c(k', k)$ has a minimum for each k and that $\min_{k'} c(k', k)$ is nonincreasing in k with $k^*(k) = \arg\min_{k'} c(k', k) \leq k$, we have to follow a different strategy. Note first, that the submodularity of $c(k', k)$ implies that $k^*(k)$ is increasing in k (Topkis (1978) , Theorem 6.1). Furthermore, submodularity together with the fact that $\min_{k'} c(k', k)$ is nonincreasing in k implies that for each (k', k) and each $\widehat{k} > k$, there exists a $\widehat{k'} \geq k'$ such that $c(k', k) = c(\widehat{k'}, \widehat{k})$.

So an increase in k increases the current period profit function and we can find an increase in k' that leaves the cost of capital investment unchanged. Provided the expectation term is nondecreasing in k' this implies that $V_S(\omega, k)$ is nondecreasing in k. Monotonicity of the expectation term in k' requires, of course, that $V(\omega, k)$ be nondecreasing in k in the first place. This can be show by an induction proof, similar to the proof of supermodularity below, where $V(\omega, k)$ is the limit of a T period horizon problem and where we have just outlined the inductive step.

3. Uniqueness: *The operator defined in equation (2.1) satisfies Blackwell's sufficient conditions for a contraction mapping as $V(\omega, k)$ is bounded and the operator satisfies the monotonicity and the discounting property as $\beta < 1$. This implies uniqueness (e.g. Stokey, Lucas(Jr) & Prescott (1989)).*

4. Supermodularity: *The proof establishes by induction the supermodularity of $V^T(\omega, k)$ in a finite horizon problem with T remaining periods and then takes the limit as $T \to \infty$.*

For the initial condition of the inductive argument simply note that after the last period ($T = 1$), the firm's continuation value is zero, i.e. $V^0(\omega, k) \equiv \max(\Phi, 0)$ which is a constant and hence supermodular.

For the inductive step denote the value conditional on survival with T periods to go

$$V_S^T(\omega, k) = \sup_{k', \psi'} \left[\pi(\omega, k) - c(k', k) - r(\psi', \omega) + \beta \int V^{T-1}(\omega', k') dF(\omega'|\psi') \right].$$

Supermodularity is preserved by optimization provided the objective function is supermodular in all its arguments (ω, k, k', ψ') (e.g. Topkis (1998), Theorem 2.7.6). Since supermodularity is also preserved under addition it is sufficient to check that each term in the brackets above is supermodular. The first three are supermodular by assumption. Supermodularity of the expectation (which is a function of (k', ψ')) follows from the definition of first order stochastic dominance and the fact that $V^{T-1}(\omega', k')$ is supermodular: Since K, Ψ, Ω are all totally ordered sets, supermodularity is equivalent to increasing differences (Topkis (1978) , theorem 3.2), or, *for any* $k_1' > k_2'$ *and* $\psi_1' > \psi_2'$,

$$\begin{aligned} & \int V^{T-1}(\omega', k_1')dF(\omega'|\psi_1') - \int V^{T-1}(\omega', k_2')dF(\omega'|\psi_1') \\ & - \int V^{T-1}(\omega', k_1')dF(\omega'|\psi_2') + \int V^{T-1}(\omega', k_2')dF(\omega'|\psi_2') \\ = & \int [V^{T-1}(\omega', k_1') - V^{T-1}(\omega', k_2')][dF(\omega'|\psi_1') - dF(\omega'|\psi_2')] \geq 0. \end{aligned}$$

The inequality holds since the difference in the value functions in the last line is an increasing function in ω *(by supermodularity of* $V^{T-1}(\omega', k')$ *which is the hypothesis of the inductive argument) and since* $F(\omega'|\psi')$ *is increasing in* ψ' *in the first order stochastic dominance sense. Therefore, the expectation term and, as a result,* $V_S^T(\omega, k)$ *is supermodular.*
Next, we show that $V^T(\omega, k) = \max[\Phi, V_S^T(\omega, k)]$ *is supermodular. Consider any* $\omega_1 > \omega_2$ *and* $k_1 > k_2$*. Supermodularity of* $V^T(\omega, k)$ *is equivalent to*

$$V^T(\omega_1, k_1) - V^T(\omega_2, k_1) \geq V^T(\omega_1, k_2) - V^T(\omega_2, k_2).$$

This inequality trivially follows from the supermodularity of $V_S^T(\omega, k)$ *if all the four terms are in the continuation region. But it also holds when some or all of the terms are in the exit region. This is because* $V_S^T(\omega, k)$ *is increasing in both its arguments so that*

$$V_S^T(\omega_1, k_1) \geq V_S^T(\omega_1, k_2), V_S^T(\omega_2, k_1) \geq V_S^T(\omega_2, k_2).$$

So whenever (ω_2, k_1) *is in the stopping region, so will be* (ω_2, k_2)*, and whenever* (ω_1, k_1) *is in the stopping region, so will be all the other terms (in which case the inequality will be an equality). As a result,* $V^T(\omega, k)$ *is supermodular which completes the inductive step.*
Finally, as $V^T(\omega, k)$ *is supermodular for every* T*,* $V(\omega, k) = \lim_{T\to\infty} V^T(\omega, k)$ *is supermodular. This completes the proof of supermodularity.* ■

Lemma 2.2 *The optimal capital choice conditional on* ψ' *and* k *of the model in equation (2.1),* $\kappa(\psi', k) = \arg\sup_{k'} \left[-c(k', k) + \beta \int V(\omega', k')dF(\omega'|\psi')\right]$,

1. *is nondecreasing in ψ' under (A0) to (A3), and*

2. *nondecreasing in k under (A0),(A1), (A3.b).*

***Proof.* 1. Nondecreasing in ψ':** *For any given k, the optimal choice k' will be nondecreasing in ψ', provided the objective function is supermodular in (k', ψ') (Topkis (1978) , Theorem 6.1). The integral is supermodular in (k', ψ') because $V(\omega', k')$ is supermodular and $F(\omega'|\psi')$ is stochastically increasing in ψ', so that the result follows. [Athey (1995), (corollary 1, example 2) also shows that* $\arg\max_{k'} \int h(\omega', k') dF(\omega'|\psi')$ *is increasing in ψ' for supermodular $h()$ if and only if $F(\omega'|\psi')$ is ordered by ψ' in the first order stochastic dominance sense.]*

2. Nondecreasing in k': *For ψ' fixed, $\kappa(\psi', k)$ is increasing in k as $-c(k', k)$ is supermodular implying that the objective function is supermodular in (k', k).* ■

Lemma 2.3 *The policy function in equation (2.4) for the* ***choice of distribution***

$$\widetilde{\psi}(\omega, k) = \arg\sup_{\psi'} \left[\pi(\omega, k) - c(\kappa(\psi', k), k) - r(\psi', \omega) + \beta \int V(\omega', \kappa(\psi', k)) dF(\omega'|\psi') \right]$$

1. *is nondecreasing in ω, under (A0),(A1), (A3.c), and*

2. *strictly increasing in ω on the set $\{(\omega, k) | r(\widetilde{\psi}(\omega, k), \omega) > 0\}$, under (A0) to (A4).*

***Proof.* 1. Nondecreasing:** *Again, by supermodularity of the objective function in (ψ', ω). The relevant term is $-r(\psi', \omega)$ which is supermodular by assumption.*

2. Strictly increasing: *The proof of strict monotonicity uses an Euler equation $G(\omega, k, \psi') = 0$ for a perturbation of the optimal ψ' between periods t and $t+1$. We will show that the $G(\omega, k, \psi')$ is strictly increasing in ω which implies that the optimal k' has to change for the Euler equation to remain satisfied. Together with part (1) this will imply the result.*
The difficulty in constructing an Euler equation for a perturbation in ψ' lies in the fact that ψ' affects the stochastic evolution of ω'. This complicates the construction of an alternative programme that leaves the joint distribution of state variables

from periods $t+2$ onwards unchanged (conditional on the state in t). We will achieve this by associating each realisation of ω' under the optimal programme with a corresponding realisation ω^* under the alternative programme. Let ψ' denote the choice of distribution under the optimal programme and consider the perturbation $\psi^* = \psi' - \varepsilon$. The next period's state variable under this perturbation, will have the distribution $F(\cdot|\psi' - \varepsilon)$. Define

$$\begin{aligned} \omega^* &= F^{-1}[F(\omega'|\psi')|\psi' - \varepsilon)] \equiv g(\omega', \psi', \varepsilon), \text{ and} \\ \Delta(\omega', \psi', \varepsilon) &\equiv \omega' - \omega^* = \omega' - g(\omega', \psi', \varepsilon). \end{aligned}$$

The function $\Delta(\omega', \psi', \varepsilon)$ has the property that $\Delta(\omega', \psi', 0) = 0$, and is differentiable as F and F^{-1} are differentiable. Using this definition, we can now formulate an alternative programme that only affects the pay-offs in periods t and $t+1$ and leaves the exit decision and the distribution of the state variables from period $t+2$ onwards unchanged. The value at period t of the alternative programme as a function of ε is then

$$\begin{aligned} V^*(\omega, k, \varepsilon) &= \pi(\omega, k) - c(k', k) - r(\psi' - \varepsilon, \omega) + \beta \int \chi(\omega', k') \{\pi(\omega' - \Delta(\omega', \psi', \varepsilon), k') \\ &\quad - c(k'', k), k') - r(\psi'', \omega' - \Delta(\omega', \psi', \varepsilon)) + \beta E[V(\omega'', k'')|\psi'', k'']\} dF(\omega'|\psi') \\ &\quad + \beta\Phi \int [1 - \chi(\omega', k')] dF(\omega'|\psi'), \end{aligned}$$

where (ω'', k'') denotes the state at $t+2$, and we use the abbreviations $\psi' = \widetilde{\psi}(\omega, k)$, $k' = \widetilde{k}(\omega, k)$, $k'' = \widetilde{k}(\omega', k')$ and $\psi'' = \widetilde{\psi}(\omega', k')$. The first integral in the expression above denotes the expected continuation value of the alternative programme in period $t+1$ when continuation is optimal under the original programme. Note that we are integrating over the distribution of ω' under the original programme and have only adjusted the arguments of the profit and R&D function using the definition of $\Delta(\omega', \psi', \varepsilon)$ above.

The difference in period t between the value function of the original and the *alternative programme is*

$$\begin{aligned} V(\omega,k) - V^*(\omega,k,\varepsilon) &= -r(\psi',\omega) + r(\psi'-\varepsilon,\omega) \\ &\quad +\beta \int \chi(\omega',k') \left\{\pi(\omega',k') - \pi(\omega' - \Delta(\omega',\psi',\varepsilon),k') \right. \\ &\quad \left. -r(\psi'',\omega') + r(\psi'',\omega' - \Delta(\omega',\psi',\varepsilon))\right\} dF(\omega'|\psi') \end{aligned}$$

By the optimality of the original programme, this expression must be non-negative in a neighbourhood of $\varepsilon = 0$ and zero at $\varepsilon = 0$. Provided differentiability, its derivative with respect to ε at $\varepsilon = 0$ must therefore be zero. This yields the Euler equation

$$G(\omega,k,\psi') = -\frac{\partial r(\psi',\omega)}{\partial \psi'} + \beta \int \chi(\omega',k')\left\{\frac{\partial \pi(\omega',k')}{\partial \omega'} - \frac{\partial r(\psi'',\omega')}{\partial \omega'}\right\}\frac{\partial \Delta(\omega',\psi',0)}{\partial \varepsilon} dF(\omega'|\psi') = 0,$$

The assumptions on $r(\psi',\omega)$ ensure that $G(\omega,k,\psi')$ is a continuous, strictly increasing function of ω for every (k,ψ'). For fixed k, and increase in ω therefore has to trigger a change in ψ' for $G(\omega,k,\psi') = 0$ to remain satisfied (recall that $k' = \kappa(\psi',k)$). Since ψ' is nondecreasing in ω by part (1) of the lemma, ψ' must be strictly increasing in ω. As the R&D function is assumed to be differentiable on the set where $r() > 0$ the result is restricted to this set. ■

Theorem 2.4 *The policy function for the* **capital choice** $\widetilde{k}(\omega,k) = \kappa(\widetilde{\psi}(\omega,k),k)$ *is*

1. *nondecreasing in ω under (A0) to (A3), and*
2. *strictly increasing in ω on the set $\{(\omega,k)|\widetilde{k}(\omega,k) > k \wedge r(\widetilde{\psi}(\omega,k),\omega) > 0\}$ under (A0) to (A4).*

Proof. 1. Nondecreasing: *The result follows directly from lemma 2.2 and lemma 2.3 as $\widetilde{\psi}(\omega,k)$ is nondecreasing in ω and $\kappa(\psi',k)$ is nondecreasing in ψ.*

2. Strictly increasing: *Again, the proof of strict monotonicity uses an Euler equation $G(\omega,k,k') = 0$ for a perturbation of the optimal capital choice k' between periods t and $t+1$. Consider the alternative programme in the continuation region: $k_t^*(\omega,k) = \widetilde{k}(\omega,k) - \varepsilon$, $k_{t+1}^*(\omega,k) = \widetilde{k}(\omega,k+\varepsilon)$, $\psi_{t+1}^*(\omega,k) = \widetilde{\psi}(\omega,k+\varepsilon)$,*

$\chi^*_{t+1}(\omega,k) = \widetilde{\chi}(\omega,k+\varepsilon)$. *The policy functions of the alternative programme are indicated by* * *and are indexed by the time period at which they apply. In all time periods other than the ones listed above, the alternative policy functions are identical to the policy functions of the original programme. The value at period* t *of the alternative programme as a function of* ε *is then*

$$\begin{aligned}
V^*(\omega,k,\varepsilon) \;=\;& \pi(\omega,k) - c(k'-\varepsilon,k) - r(\psi',\omega) \\
& +\beta \int \chi(\omega',k')\Big\{\pi(\omega',k'-\varepsilon) - c(\widetilde{k}(\omega',k'),k'-\varepsilon) - r(\widetilde{\psi}(\omega',k'),\omega') \\
& \qquad +\beta E[V(\omega'',\widetilde{k}(\omega',k'))|\widetilde{\psi}(\omega',k')]\Big\}\, dF(\omega'|\psi') \\
& +\beta\Phi \int [1-\chi(\omega',k')] dF(\omega'|\psi'),
\end{aligned}$$

where k' *and* ψ' *denote the optimal policies at time* t *under the original programme. The second and third line of the expression above denote the value of the alternative programme when continuation is optimal under the original programme.*

By the optimality of the original programme, the difference between the value of the original and the alternative programme be non-negative in a neighbourhood of $\varepsilon = 0$ *and zero at* $\varepsilon = 0$. *Provided differentiability, its derivative at* $\varepsilon = 0$ *must therefore be zero. This yields the Euler equation*

$$G(\omega,k,k') = -\frac{\partial c(k',k)}{\partial k'} + \beta \int \chi(\omega',k')\{\frac{\partial \pi(\omega',k')}{\partial k} - \frac{\partial c(\widetilde{k}(\omega',k'),k')}{\partial k}\} dF(\omega'|\widetilde{\psi}(\omega,k)) = 0.$$

The assumptions and the form of the optimal policy ensure that $G(\omega,k,k')$ *is a continuous function of* ω *for every* (k,k'). *Furthermore,* $\chi(\omega',k')$ *is non-decreasing in* ω' *(see equation (2.2) and the discussion thereafter),* $\frac{\partial}{\partial k}\pi(\omega',k')$ *is strictly increasing in* ω' *and* $\frac{\partial}{\partial k}c(k',k)$ *is non-increasing in* k' *by assumption and* $\widetilde{k}(\omega',k')$ *is non-decreasing in* ω' *by part (1) of the proof. This implies that the integrand is strictly increasing at some* ω' *provided continuation is optimal for some* ω'. *This latter point has to be true if* $k' > k$, *as not increasing the capital stock in period* t *would reduce the cost of investment and yield a strictly higher value if exit in* $t+1$ *was certain. So the integrand is non-decreasing* ω' *and strictly increasing in* ω' *at some* ω' *within the domain of* $F(\omega'|\widetilde{\psi}(\omega,k))$.

As $\widetilde{\psi}(\omega,k)$ *is strictly increasing in* ω *on the set where* $r > 0$ *by Lemma 2.3, and*

$F()$ is strictly stochastically increasing in ψ', the integral in the expression above is increasing in ω. With k fixed, this implies that k' needs to adjust for the Euler equation to remain satisfied. Together with the weak monotonicity of part (1) this completes the proof. ■

A2.2: Data and construction of variables

Selection of Industries

The four industries in this study are selected on the basis of the following criteria: (1) All 3-digit SIC industries with fewer than 3000 firm year observations over the period 1980-2001 are deleted, where each firm is allocated to the 4-digit industry which it reports as its primary industry of operation. (2) As the focus is on the effect of R&D on the distribution of productivity, industries with fewer than 1000 observations on R&D expenditure are dropped. (3) Of the remaining 3-digit SICs "Measurement Instruments (SIC 382)", "Medical Apparatus (SIC 384)", and "Electronic Equipment (SIC 367)" are dropped completely, as there seems to be a wide variation in the nature of the constituent 4-digit sub-industries.

This leaves the four 3-digit industries in this study. To construct relatively homogenous industries, we further drop some 4-digit industries from the remaining set of 3-digit industries which we feel differ significantly from the other 4-digit industries falling in the same 3-digit group: From "Pharmaceuticals", SIC 2833 "Medicinal Chemicals and Botanical Products" (181 obs.) is excluded. From "Computer Hardware", we exclude firms coded under SIC 3570 "Computer and Office Equipment" (171 obs.), SIC 3576 "Computer Communication Equipment" (1224 obs.), SIC 3578 "Calculating and Accounting Machines excluding Computers" (358 obs.), and "SIC 3579 Office Machines, not elsewhere classified" (236 obs.). From "Telecommunications Equipment" we exclude SIC 3669 "Communications Equipment, not elsewhere classified" (636 obs.). From "Software" we exclude SIC 7370 "Computer Programming, Data Processing" (2909 obs.), SIC 7374 "Computer Processing, Data Preparation Services" (559 obs.), and SIC 7377 "Computer Rental and Leasing" (154 obs.). Table 2.1 reports the remaining number of firms and firm years in the estimation sample by 3-digit industry and constituent 4-digit industry.

Capital stock

The capital stock is constructed using the standard perpetual inventory method where we set the capital stock of a firm's second year in the data (the initial capital stock for our purposes) equal to the firms book value of the net stock of property, plant, and equipment in the COMPUSTAT database (CS data item A8 – PPENT) in its first year of the data (i.e. the book value in the opening balance of the second year). We deflate the initial capital stock using the price deflator for investments in that year which is described below. The capital stock in subsequent years is then constructed as the depreciated capital stock in the previous year plus capital investments (CS data item A39 – CAPXV). Again, investments are deflated as described in below. Since we have set the "initial capital stock" for the firm's second year, we also construct the capital stock in the first year using the perpetual inventory method.

Some companies report rental expenditure for renting equipment and space (CS data item 46 – XRENT). Once the capital stock is constructed, we capitalise these rental expenditure using the depreciation rate plus a rental premium of 2% and add them to the capital stock.

The choice of depreciation rates for the capital stock is a delicate issue. Many previous studies (among them Olley & Pakes (1996)) employed the depreciations rates for equipment and structures reported by Hulten & Wykoff (1981). We think this is problematic in our case for two reasons: first, we do not have detailed information for each firm on the types of equipment and structures in place. For example, Hulten & Wykoff (1981) report a depreciation rate for "communications equipment" of .1179 and for "office, computing, and accounting machinery" a rate of .2729. Second, the study employs data from the 1970s and we feel that the rapid technological progress in information technology and other fields may have lead to an acceleration of the obsolescence of equipment. However, Hulten & Wykoff (1981) also report depreciation rates for equipment and structures for which they only have length of life estimates. To do this they employ the relationship $\delta = R/L$, where L is the length of life estimate and R is the "declining balance rate". They estimate R separately for equipment and structures on the subsample of asset categories on which they have detailed data. Their point estimate is 1.65 for equipment and .91 for structures. We will employ these point estimates to construct our depreciation rates. Using a similar approach as Hall (1990), we estimate the length of life of assets in each industry as the median across firms and time of the ratio between gross property plant and equipment to depreciation (CS data item A7 – PPEGT divided by CS data item A14 – DP). We report this median length of life estimate in the second column in the table below. To obtain an approximation of the percentage of investment that goes, on average, towards equipment we use the ratio of the change in the gross stock in materials and equipment over the change in the total gross property, plant, and equipment (the change in CS data item A264 – PPECME divided by the change in CS data item A7 – PPEGT). The resulting "weights" are reported in the third column in the table below. We then calculate our estimate for the depreciation rate δ as $[1.65w + .91(1 - w)]/L$ and report it in column four of the table. The next two columns report the corresponding "single-declining balance depreciation rate" δ_1 and "double-declining balance depreciation rate" δ_2 for comparison. Finally, we report the implied depreciation rates for aggregate stocks in the "NBER-CES Manufacturing Industry Database (1958 - 1996)" over the period 1980-1997 using the perpetual inventory formula. The implied NBER depreciation rates are far below our estimates. We feel that our estimates δ in column four are more reasonable and use them for the construction of the capital stock.

Industry	L	w	δ	δ_1	δ_2	$NBER$
283	8.6	.76	.1710	.1161	.2323	.0427
357	6.6	.92	.2415	.1522	.3043	.0579
366	7.8	.88	.2003	.1281	.2563	.0572
737	5.5	.89	.2843	.1816	.3632	–

The price deflator for investments for 1980-1996 is taken from the NBER database. It is constructed from input-output tables and deflators for 28 types of asset classes (Bartelsman & Gray (1996)). To extend this series until 2001 we take the following approach. A price index for investment is a weighted average of price deflators for investments in property, plant and equipment. An inspection of the

BEA's table on investments in nonresidential fixed assets by two digit industry and detailed asset type (http://www.bea.gov/bea/dn/faweb/Details/Index.html) confirms that since 1987 the main investments in the industries under study fall into the categories of computer hard and software, communications equipment, industrial machinery and industrial buildings. So we try to fit a composite deflator of these underlying deflators to the NBER investment deflator and then use the predictions for 1997-2001 as our investment deflator for this period. This approach is, of course, problematic: First, for each 4-digit industry, we only have ten data points (1987-1996) for our regressions of the NBER deflator on the underlying deflators. For this reason, we choose a very parsimonious specification, in simply regressing the NBER deflator on one price deflator for construction industries in SIC's 15 to 17 (obtained from the BEA table "1947-2001 Gross Domestic Product by Industry NDN-0302 ") and on the deflator for "industrial machinery and equipment – SIC 35" (from the BEA tables on "1977-2001 Manufacturing Industry Shipments NDN-0304"). We choose these two deflators as regressors, because SIC 35 includes computers. Furthermore, it is very highly correlated with the deflator for "Electronic and other electric equipment – SIC 36" so that one deflator for "equipment" should suffice. Including a price deflator for software did not seem to improve the fit dramatically. So for each of our 4-digit manufacturing industries, we run no intercept OLS of the NBER investment deflator on the price deflator for construction industries and the deflator for industrial machinery. We then use the estimated "weights" to predict the investment deflator for 1997-2001. The second problem with this approach is the fact that we are making out of sample predictions and that the weights are likely to change over time. To investigate the quality of the predictions, we run additional regressions where we only use the first (respectively the last) five years in the ten year sample. While the predictions for 1992-1996 from the regressions on 1987-1991 do not seem to fit the period 1992-1996 very well, the backward predictions from the regressions on the subsamples for 1992-1996 perform much better. Furthermore, the predictions for 1997-2001 from the second subsample are very close to the predictions based on the whole sample. Finally, the predictions for 1997-2001 based on the whole sample seem to continue the trend from the NBER deflator series 1987-1996 reasonably well. Together, these facts give us some confidence that the predicted values for the investment deflators for 1997-2001 are not completely wrong. The lack of an obviously superior deflator leads us to employ these predictions in the construction of our capital stock.

Lacking a similar price deflator for the software industries (SIC 737), we use the GDP deflator to deflate investments.

Value added

We construct value added as a measure of output minus cost of materials. We measure output as sales revenue (CS item A12 – SALE) plus the change in the inventory of finished goods (CS item A78 – INVFG). There are two items in the COMPUSTAT database covering operating costs: "Costs of goods sold" (CS item A41 – COGS) and "selling, general, and administrative expense" (CS item A189 – XSGA). These items include costs of materials, but also R&D expenditure (CS item A46 – XRD), expenditure for the rental of equipment and space (CS item A46 – XRENT), and labour costs. To obtain a measure of costs of materials, we subtract R&D and rental expenditure from the measure of operating costs (rental expenditure is capitalised and added to the capital stock).

Two data items on labour costs are available which are also included in the variable cost measure: "Labour and related expense" (CS data item A42 – XLR) and "Pension and related expense" (CS data item A43 – XPR). Unfortunately these items are missing in many cases. Only just over one half of the observations on pension expenses in non-missing and only about 5% of the labour expense data is available. Obviously, this is a very severe data restriction. For each 3-digit industry, we impute measures of pension and labour expenses for all companies using the following procedure: The first step is to deflate "pension and related expense" and "labour and related expense" by a wage deflator constructed from BEA data on total employee compensation for total manufacturing and the full time equivalent of workers. We then run OLS regressions of these deflated variables log employment, log capital, log employment squared and capital squared. This specification was chosen because it seemed to yield a satisfactory fit of the predicted values with the actual data, without obviously over or underpredicting expenditure of big or small companies. We then subtract inflated predicted values from these regressions from our costs of material measure. Value added is then constructed as gross output minus cost of materials and deflated by an output deflator.

The output deflator is constructed as follows: For all manufacturing industries, the output price deflator for the period 1980-1997 is the price index for the value of shipments from the "NBER-CES Manufacturing Industry Database (1958 - 1996)" on the basis of the 1987 SIC codes. To extend this series until 2001, we turn to the US Department of Commerce Bureau of Economic Analysis (BEA) tables on "1977-2001 Manufacturing Industry Shipments NDN-0304". The BEA table for 1987-2001 uses the same industry definitions as the NBER database (1987 SIC codes). For each manufacturing 4-digit industry in our study, we compare the implicit price deflator for the value of shipments from the BEA table for 1987-1997 with the corresponding NBER deflator. Over this period, the two deflator series are almost identical. Only for SIC 3661, one can visually detect a small gap in the indices over this time period. However, even this gap is small compared to the gap between the NBER productivity deflator and other available deflators (such as the Bureau of Labor Statistics' Producer Price Index). Therefore, we employ the BEA deflator for manufacturing industries for the period 1997-2001.

For the software industries (SIC 737), we lack a consistent price deflator over the entire period and we simply employ the US GDP deflator.

Labour

The labour input is the number of employees (CS data item A29 – EMP).

R&D

We use R&D expenditure as reported by the firm (CS data item A46 – XRD) deflated by the output price deflator for the industry.

Other variables used

In the estimation, we will exclude observations that have grown very rapidly due to mergers or acquisitions. The item we use to determine these events "Sales contribution of acquisitions" (CS data item A249 – AQS).

Bibliography

Athey, S. (1995), 'Characterizing properties of stochastic objective functions', *mimeo* .

Bartelsman, E. J. & Gray, W. (1996), 'The NBER manufacturing productivity database', *NBER Technical Working Paper 205* .

Blundell, R. & Bond, S. (2000), 'GMM estimation with persistent panel data: An application to production functions', *Econometric Reviews* **19**(3), 321–340.

Buchinsky, M. (1994), 'Changes in the U.S. wage structure 1963-1987: Application of quantile regression', *Econometrica* **62**(2), 405–458.

Cox, D. & Miller, H. (1965), *The Theory of Stochastic Processes*, William Clowes and Sons, London.

Dixit, A. K. & Pindyck, R. S. (1994), *Investment under Uncertainty*, Princeton University Press, Princeton.

Ericson, R. & Pakes, A. (1995), 'Markov-perfect industry dynamics: A framework for empirical work', *Review of Economic Studies* **62**, 53–82.

Griliches, Z. (1998), *R&D and Productivity*, The University of Chicago Press, Chicago.

Hall, B. H. (1990), 'The manufacturing sector master file: 1959-1987', *NBER Working Paper No. 3366* .

Hopenhayn, H. A. (1992), 'Entry, exit, and firm dynamics in long run equilibrium', *Econometrica* **60**(5), 1127–1160.

Hulten, C. & Wykoff, F. (1981), The measurement of economic depreciation, *in* C. Hulten, ed., 'Depreciation, Inflation, and Taxation of Income from Capital', Urban Institute Press, Washington D.C.

Jovanovic, B. (1982), 'Selection and the evolution of industry', *Econometrica* **50**(5), 649–70.

Koenker, R. & Basset(Jr), G. (1978), 'Regression quantiles', *Econometrica* **46**(1), 33–50.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *Review of Economic Studies* **70**, 317–341.

Mairesse, J. & Sassenou, M. (1991), 'R&D and productivity: A survey of econometric studies at the firm level', *NBER Working Paper No. 3666* .

Mansfield, E. (1962), 'Entry, gibrat's law, innovation, and the growth of firms', *American Economic Review* **52**(5), 1023–1051.

Marschak, J. & Andrews, W. H. (1944), 'Random simultaneous equations and the theory of production', *Econometrica* **12**, 143–205.

McGahan, A. M. (1999), 'Competition, strategy, and business performance', *California Management Review* **41**(3), 71–104.

Milgrom, P. & Shannon, C. (1994), 'Monotone comparative statics', *Econometrica* **62**(1), 157–180.

Mueller, D. C. (1986), *Profits in the Long Run*, Cambridge University Press, London, New York.

Mueller, D. C., ed. (1990), *The Dynamics of Company Profits - An International Comparison*, Cambridge University Press, Cabridge and New York.

Olley, S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.

Pakes, A. (1994), Dynamic structural models, problems and prospects: Mixed continuous discrete controls and market interactions, *in* C. Sims & J. Laffont, eds, 'Proceedings of the 6th World Congress of the Econometric Society, Barcelona, Spain', Cambridge University Press.

Pakes, A. & Olley, S. (1995), 'A limit theorem for a smooth class of semiparametric estimators', *Journal of Econometrics* **65**, 295–332.

Robinson, P. M. (1988), 'Root-N-consistent semiparametric regression', *Econometrica* **56**(4), 931–954.

Stokey, N. L., Lucas(Jr), R. E. & Prescott, E. C. (1989), *Recursive Methods in Economic Dynamics*, Havard University Press, Cambrige, Massachusetts.

Sutton, J. (1998), *Technology and Market Structure*, The MIT Press, Cambridge MA, London England.

Sutton, J. (2002), 'The variance of firm growth rates: The scaling puzzle', *Physica A* **312**, 577–590.

Topkis, D. M. (1978), 'Minimizing a submodular function on a lattice', *Operations Research* **26**(2), 305–321.

Topkis, D. M. (1998), *Supermodularity and Complementarity*, Princeton University Press, Princeton, New Jersey.

Wedervang, F. (1965), *Development of a Population of Industrial Firms*, Scandinavian University Books, Oslo, Norway.