CPB Document

No 180

March, 2009

Relation entry, exit and productivity

An overview of recent theoretical and empirical literature

Viktoria Kocsis, Ruslan Lukach, Bert Minne, Victoria Shestalova, Nick Zubanov and Henry van der Wiel

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN 978-90-5833-396-4

Abstract in English

This document provides a review of recent theoretical and empirical literature on the relationship between entry, exit and productivity. Decomposition methods show that entry and exit considerably contribute to productivity growth, but are unable to shed any light on the ultimate sources of productivity growth. However, the theories discussed do provide options for effective policy instruments. We argue that productivity or welfare should be the aim of policy and not the number of entrants, the intensity of competition or the amount of innovation expenditures. Taking a welfare approach, we address market failures with respect to entry. The most eminent market failure is market power of dominant incumbents. Lowering institutional entry barriers economy-wide is a promising policy option for further consideration. Whether such a policy measure actually improves social welfare depends also on the extent of other failures. Therefore, an ex ante cost-benefit analysis needs to precede intervention.

Key words: entry, exit and productivity
JEL code: B41, O30, 040

Abstract in Dutch

Dit document geeft een overzicht van de recente theoretische en empirische literatuur over de relatie tussen toetreding, uittreding en productiviteit. Decompositiemethoden laten zien dat toe- en uittreding sterk bijdragen aan de toename van de productiviteit. Deze methoden geven echter geen zicht op de oorzaken van productiviteitsgroei. De besproken theorieën geven die inzichten wel en laten daarmee zien wat de mogelijke opties zijn voor effectieve overheidsinstrumenten. We stellen dat productiviteit of welvaart het doel voor beleid zou moeten zijn en niet het aantal toetreders, de mate van concurrentie, of de -uitgaven aan innovatie. Deze laatste zijn alleen maar middelen om productiviteit te genereren. Uitgaande van een welvaartsbenadering adresseren we marktimperfecties gerelateerd aan toetreding. De belangrijkste marktimperfectie is marktmacht van dominante bedrijven. Het wegnemen van institutionele toetredingsbelemmeringen is een nader te overwegen beleidsoptie. Of deze optie ook leidt tot meer welvaart, hangt af van andere marktimperfecties. Hiervoor is eerst een kosten-batenanalyse vereist.

Steekwoorden: toe- en uittreding, productiviteit

Een uitgebreide Nederlandse samenvatting is beschikbaar via www.cpb.nl.

Contents

Preface

New firms are often considered as a key source for an economy, particularly for stimulating economic growth and productivity. For that reason, policy puts effort in increasing the number of entrepreneurs.

This document provides insights into the relationship between entry, exit, innovation and productivity. It discusses the empirical significance of these key concepts and presents the current theoretical views on their relationship to each other, stressing the direction of causality. The knowledge of causality is essential for designing effective policies to promote productivity of the economy. This document derives implications for entry policies.

Viki Kocsis, Victoria Shestalova, Ruslan Lukach, Bert Minne, Nick Zubanov and Henry van der Wiel (project leader) wrote this document. We express our thanks to our CPB colleagues Paul de Bijl, George Gelauff, and Free Huizinga for their comments on earlier versions of this document. We also thank Hugo Erken and Willem Spruijt (both Ministry of Economic Affairs).

Coen Teulings
Director CPB

Summary

Research questions

Productivity growth is one of the key goals of economic policy. Policy makers regard entrepreneurship as an important mechanism for achieving this goal as they link it with innovation and sound risk taking. Their idea is that new firms, or entrants, in the jargon of the literature, embody the innovations. Therefore policy makers promote and stimulate entry.

To be effective, growth-enhancing policies need specific knowledge about what causes what, because only by affecting the sources which cause productivity to grow their goal may be achieved. To help policy makers design effective policies towards entrepreneurship, this document investigates the following questions. First, what are the stylised facts on the relationships among entry, exit, innovation and productivity? Second, what do theories tell about the mechanisms behind these relationships? Finally, we derive from the answers to both questions the implications for Dutch policy makers to design effective policies regarding entrepreneurship.

Main conclusions

Empirical decomposition studies of productivity growth show that entry and exit of firms contribute to aggregate productivity growth considerably. The econometric evidence, however, provides an ambiguous correlation between business ownership and productivity growth. Both types of studies do not say anything about the causality between entry, exit and productivity. Theory can say more about this causality and the general conclusions from our review are as follows. While entry, actual or the threat of entry, does encourage at least some of the incumbents to become more efficient, entry itself is driven by market opportunities arising either exogenously or as a result of entrants' own efforts (e.g., innovation activities). High as well as low entry rates can go together with productivity growth. Therefore, the message from theories to policy makers is that entry is only a mechanism through which other factors contribute to productivity growth. Entry per se should not be the goal of economic policies, but affecting the size of entry costs can be a potential instrument to improve productivity through competition and/or innovation. Reflecting a number of options related to those costs, reducing institutional entry barriers seems to be a promising policy measure. Whether such a measure actually improves social welfare depends also on the extent of other market failures. Therefore, an ex ante cost-benefit analysis needs to precede intervention taking alternative options into account that can stimulate productivity through entry as well.

Summary of the document

We begin with an introduction in chapter 2 of the key theoretical concepts used in the remaining of the document: entrepreneurship, firm entry and exit, entry and exit costs and barriers,

productivity, efficiency and welfare. Entry and exit of firms are key mechanisms that deliver new ideas to the market, hence their importance. Entry costs are expenses only paid by new start-ups (licenses, permits etc.), and are therefore often labelled as barriers to entry. The costs of exit include asset divestment costs, such as severance pay to laid-off workers. The higher the costs the less intensive entry and exit are, but not all of these costs are legitimate targets for policy intervention as we discuss later. In general, high productivity corresponds to high welfare – being the present value of the sum of consumer and producer surplus – since GDP per capita increases if productivity rises. But a trade-off might exist between static and dynamic efficiency. The former reflects changes in aggregate productivity performance under a given technology in the short-term. The latter corresponds to long-term productivity improvements due to technological changes or discovery of new products.

We then present in chapter 3 a selection of stylised facts on entry and exit. A large share of firms (around 20%) enters and exits a particular market each year. The size and productivity level of entrants is lower than that of incumbents, and their 'mortality rate' is high in the first years of their existence, but the surviving ones grow faster both in size and productivity than incumbents. Later we present some theories, in particular, active learning models, explaining this empirical regularity. Entry and exit are positively correlated most of the times, but the correlation between the two, as well as the intensity of entry and exit, varies by country. As a potential determinant of this variation, we propose differences in institutional factors affecting entry and exit.

We then report findings of the contribution of entry and exit to aggregate productivity growth. Empirical decomposition studies of productivity growth show that entry and exit of firms contribute to aggregate productivity growth considerably. The contribution of entry and exit is both direct, through entry of more productive firms and exit of less productive firms, and indirect, through changes in the market shares of incumbents. The decomposition results vary with industry, the length of time period captured and the estimation method, but the general impression is that the direct and indirect effects of entry together are responsible for a large part (up to 70%) of the observed productivity growth. The econometric evidence, however, provides an ambiguous correlation between business ownership and productivity growth. Data availability mainly hampers econometric studies to cope with the endogeneity problem. Both the decomposition studies and the econometric studies do not say anything about causality between entry, exit and productivity. Theory can say more about this causality.

Chapter 4, therefore, presents a selection of four theoretical models of firm entry and productivity dynamics. The classical view on entry and productivity argues that the threat of entry or actual entry encourages incumbent firms to improve their efficiency in order to withstand competition. However, most of the more recent models imply that entry per se does not affect aggregate productivity. Instead entry, exit and productivity are influenced by third factors. For instance, the model of Cabral (2002) and passive learning models assume that exogenous technological changes affect productivity, and those changes drive entry.

Allowing an extra layer of complexity, active learning, life cycle and Schumpeterian growth models include the possibility of productivity growth through innovation. These models are related to the endogenous growth literature. The consequences with respect to the link between productivity and entry and exit are several. In particular, incumbent firms may affect entry by employing entry deterrence strategies. Entry and exit, and their correlation with productivity, differ in intensity depending on the stage in the industry life cycle. The threat of entry can have a positive as well as a negative effect on the incumbent's propensity to innovate, depending on how technologically advanced it is. These more advanced theories are able to explain most of the stylised facts reported in chapter 3, suggesting their consistency with the real world. But at the same time the link between entry and productivity becomes increasingly complicated and unequivocal.

The general conclusions from our review of the theories are as follows. While entry, actual or the threat of, does encourage at least some of the incumbents to become more efficient, entry itself is driven by market opportunities arising either exogenously or as a result of entrants' own efforts (e.g., innovation activities), and deterred by entry barriers. High as well as low entry rates can go together with productivity growth. Therefore, the message from theories to policy makers is that entry is only a mechanism through which other factors contribute to productivity growth. Entry per se should not be the goal of economic policies, but affecting the size of entry costs can be a potential instrument to stimulate higher productivity through competition and/or innovation.

Chapter 5 looks at policy options with respect to firm entry using a welfare economic approach. Welfare is not optimal if market failures or government failures are present. If market failures exist, government may consider intervening if and only if government failures are smaller than those market failures; i.e. if the social benefits of intervention exceed the costs. We particularly consider three potential market failures: market power, capital market failures and, finally, production externalities related to knowledge spillovers and business stealing effects. Market power of incumbents occurs if incumbents can prevent entry or if institutional barriers protect them. A capital market failure is due to asymmetric information between the supplier of funds (e.g. banks) and potential entrants. R&D or innovation activities generate knowledge spillover effects creating benefits that cannot be fully appropriated by the inventor.

For each market failure we highlight the failure, the consequences, empirical evidence and policy options taking account of the policies already in place to repair that failure. In fact, the existing competition legislation should prevent the abuse of market power of incumbents. Well-designed intellectual property rights and innovation subsidies should address production externalities directly, not regulating entry. Those instruments tackle the core problem of the market failure at its roots. Finally, capital rationing to new firms is often not a market failure.

The supply of credit may be less due to new ventures being inherently more risky, which is a fair market outcome.

Reducing (institutional) entry barriers seems to be the most promising channel for policy makers to improve productivity via entry. In that respect, it is an interesting finding that the Netherlands has no top ranking on (components of) product market regulation and on employment protection in an international perspective if countries are ranked from low to high entry barriers. This finding needs further consideration including the following issues. First, one should be careful to draw firm conclusions based on those rankings as rankings can be biased due to measurement issues. Second, the outcome can be the preferred outcome of a particular country given its choices with respect to efficiency and equity issues. Finally, government interventions are only justified as long as social benefits are larger than the costs of intervention. Therefore, an ex ante cost-benefit analysis needs to precede each intervention. Such an analysis should also consider alternative entry policies that can affect productivity as not only entry costs determine whether a firm enters the market. Examples are policies that (ex ante) reduce exit costs and labour market rigidities related to market failures. A cost-benefit analysis is, however, beyond the scope of the document.

Besides summarising the main conclusions, chapter 6 also addresses questions for a research agenda. Two topics need at least further consideration. First, as theory is ambiguous on the relationship between competition, innovation and productivity and hence the implications for policy, the evidence rests on empirical studies. With regard to the Netherlands, empirical evidence for an inverted-U link between competition and innovation is scarce. Hence, this relationship represents an interesting avenue for further empirical research based on Dutch firm-level data. Second, the empirical evidence related to labour market regulation and productivity is also scarce and provides no decisive answer to the relationship between the labour and product market. Therefore, the interaction of labour market regulation and product market regulation and its impact on productivity is another interesting topic for further research.

1 Introduction

Setting the stage

Productivity and entrepreneurship are the key words in this study. The two are interrelated. Higher productivity, essential for the growth in (real) Gross Domestic Product (GDP) per capita, can be – amongst others – achieved through stimulating entrepreneurship. It is not surprising therefore that firm creation and entrepreneurship are also the key words in many policy notes across Europe, including the Lisbon goal of 'creating a dynamic and innovative economy'.[1]

Entrepreneurship has a long history in economics. Joseph Schumpeter advocated the entrepreneur as the driving force of economic development. He called 'creative destruction' the process whereby obsolete, less efficient and therefore less competitive firms surrender to the new, more innovative and efficient entrants. In that sense, entrepreneurship and entry are interchangeable. New firms create jobs and are often considered to be the main driver of innovation. Hence, new firms are a key source for stimulating economic growth and productivity.[2]

In that vein, it is not that surprising that consecutive Dutch governments are concerned with (a lack of) entrepreneurship in the Netherlands. Therefore, they have put in efforts to increase the number and quality of entrepreneurs in the last decade.[3]

Goal and research questions document

The main goal of this document is to provide insights into the relationship between entry, exit, innovation and productivity.[4] The document discusses the significance of these key notions and presents the current theoretical views on their relationship to each other, stressing the mechanisms and the direction of causality. The knowledge of causality is essential for designing effective policies to promote productivity of the economy. In particular, the discussion in this document addresses three research questions:

- What are the stylised facts on the relationship between entry, exit, innovation and productivity?
- What do theories tell about the mechanisms behind this relationship?
- What do these theories and stylised facts imply for Dutch policy?

[1] See e.g. EC, 2003, and EC, 2004.
[2] Chapter 2 discusses both the concept of new firms and entrepreneurship in more detail.
[3] See "In actie voor ondernemers" (2003), and the recent letter of the cabinet: "werkend aan ondernemend Nederland: het belang van zelfstandig ondernemerschap (2008).
[4] Note that education and other human capital issues are also important drivers of productivity growth, but this document mainly focuses on entry, exit and innovation in relationship with productivity growth. Furthermore, we leave aside the importance of firm growth and survival.

To see how important entry and exit can be for productivity growth, the document provides in chapter 3 some stylised facts and figures for a number of countries including the Netherlands. Therein we will present the results of several decompositions of labour productivity growth into the contribution of incumbents, entrants, exits and reallocation effects. These decomposition methods link individual productivity levels to aggregated industry productivity levels. It turns out that in terms of contribution entry and exit can be substantial for productivity growth.

But, any productivity growth decomposition is only a statistical tool. It neither provides information about the driving forces behind the contributions of each component to productivity, nor takes into account the threat of entry. Hence, these tools do not give clues on the causality between entry/exit and productivity growth and the links in between (such as innovation).

Theories do provide insights into the mechanisms driving the relationship between entry, exit and productivity. Therefore, we move to the literature review of recent theories to understand this causality (see chapter 4). Theoretical results will be evaluated based on empirical papers analysing the same models. This approach is in line with existing reviews, such as the EC paper (2005), on the impact of market entry and exit on EU productivity and growth performance. We extend these reviews by incorporating the mechanisms (including the exogenous factors) underlying productivity changes. We show and explain the impact of those determinants on entry and productivity.

Indeed, from a policy perspective, it is important to know how changes in exogenous factors affect entry, productivity and firms' efforts to innovate. This determines the scope for effective policy to facilitate entry and to promote productivity and innovation. In that respect, it is also important to take into account different levels of aggregation - i.e. micro, industry and macro level - and reallocation issues within and across industries, as well as firm exit. We therefore classify the theories along potential exogenous factors to see what policies (if at all) these different theories imply.

Effective policies must trigger the right exogenous factors. Given the theories, we can determine whether entry/exit is exogenous or endogenous and address the ensuing policy options (see chapter 5). In the Netherlands, entry is likely to be largely endogenous to technological innovation coming from elsewhere in the world, given the special characteristics of the Dutch economy: small size, openness, and high productivity level.

Market failures – i.e. situations where a competitive market outcome does not deliver the maximum welfare – legitimise policy interventions if the social benefits are larger than the cost of intervention. We discuss potential market failures directly connected to entry (and exit). Those failures include part of the institutional barriers and capital market failures.

Structure of the document

The structure of this document is as follows. Chapter 2 discusses several concepts relevant for this study, such as entrepreneurship, entry and efficiency. Readers already familiar with these concepts can easily skip this chapter. Chapter 3 presents some stylised facts on entry and exit and their contribution to productivity growth. Chapter 4 reviews the mainstream literature and outlines the main drivers for entry and productivity. Next, we discuss policy implications of theories regarding entry with the focus on the Netherlands. Chapter 6 ends with concluding remarks and addresses open issues.

2 Key concepts

This document uses a number of concepts. Here, we give short definitions of the most important ones. These are: entrepreneurship, entry (costs), exit (costs), static and dynamic efficiency, productivity and welfare. Readers already familiar with these concepts can easily skip this chapter.

2.1 Entrepreneurship, entry and exit

Entrepreneurship and its link to industrial dynamics

There is no generally accepted definition of entrepreneurship, but most definitions are somehow related to the words 'new' and 'opportunity'. In the popular domain, Wikipedia defines entrepreneurship as 'the practice of starting new organizations or revitalizing mature organizations, particularly new businesses generally in response to identified opportunities'. A more official definition of entrepreneurship is given in the 'Green paper: Entrepreneurship in Europe' (EC, 2003), as the mindset and process of creating and developing economic activity by blending risk-taking, creativity and/or innovation with sound management, within a new or an existing organisation.

A survey in Verheul et al. (2002) allows us to identify the key functions of the entrepreneur. They are: perception of new economic opportunities, making decisions on resource allocation, management of business ventures, and introducing new products to the market (innovation). Both incumbent firms and new entrants may perform the latter function. In the first case, it is sometimes called 'entrepreneurship'.[5] The latter case is more in the focus of the economic theories of entrepreneurship, because it necessitates the discharge of all other functions entering the definition of this activity.

The term entrepreneurship integrates the (activities of) incumbents and industrial dynamics, i.e. entry and exit of firms. Our research relates to the issue of entrepreneurship, which is high on the current economic policy agenda in the Netherlands, in that new business ideas are delivered to the market through the arrival of new firms. Another aspect which brings entrepreneurship and firm entry and exit together is the activities of incumbents. Economic theories (e.g., Aghion et al., 2004) suggest that these activities can be entrepreneurial or counter-entrepreneurial, depending on how technologically advanced the industry in question is. We elaborate more on the theoretical interplay between incumbents and entrants in chapter 4 and derive several implications for economic policy in chapter 5.

[5] W. Baumol introduced this term.

Entry and exit as a measure of entrepreneurship

Figures on entrepreneurship and industrial dynamics are often reported to underline the importance of these concepts for economic growth and structural changes (see chapter 3). One available measure of entrepreneurial activity is business-ownership rate: the share of self-employed in the total labour force. This measure is, in a sense, rigid, because it does not quantify the creation of new business ventures, which is an important aspect of entrepreneurship. In contrast, firm entry and exit rates are more flexible measures of entrepreneurship. Those *rates* can be calculated as either the number of firms entering or exiting the market relative to the total number of firms in existence within a given period of time.[6]

2.2 Entry and exit costs

Entry and exit costs

The decision of a firm whether or not to enter the market depends on a number of (cost related) issues. In case of profit maximization, a firm or new entity decides to enter the market if its discounted total expected future profits are larger than the sum of the (sunk) entry costs and discounted exit costs.[7] If it costs a lot to start a new business, then few new businesses will start up. Hence, reducing the entry costs (given the exit costs and profits) will have a positive effect on the extent of entry. Potential entrants are more inclined to enter the market. Similarly, lowering the exit costs will have a positive effect on entry all else equal. The exit costs depend, amongst others, on the time preference of the firm. The time preference of the entrant determines the rate of discount. If the 'firm' has a high time preference, it values the present relatively higher than the further future. Hence, the higher the time preference, the higher the discount used on both benefits and costs.

What comprises entry costs? Broadly defined, entry costs include all costs a start-up must bear before it can operate legally. Entry costs include, among others, investments in plants, buildings, machinery, employees, and expenses of licenses and permits, legal and notary charges. Besides those costs, potential entrants come across non-monetary costs such as the number of procedures and permits or strategic behaviour of incumbents.

Exit costs are costs required to exit the market, such as firing employees and terminating contracts with suppliers or buyers. These costs may be an obstacle for firms to leave the market. Again, in case of rational behaviour and profit maximisation, if a firm enters the market it already takes into account the sunk costs (of exit). This changes if firms are already active in an industry. Then those costs are not anymore seen as opportunity costs of production.

[6] Note that an entry can take different forms: with or without employees, an existing company that enters a new industry, a foreign enterprise that enters the market.

[7] Note that the costs of entry and exit are additive and not multiplicative in the firm's decision to enter the market. Moreover taxes, subsidies and adjustments costs during operation have an impact on this decision as well.

Difference between private versus social entry and exit costs

In principle, the costs of entry and exit are borne by entrants or firms that exit the market. However, in some cases the social costs of entry and exit can be larger than private costs. In these cases, there are external costs related to negative externalities. Examples of external costs are costs caused by the business stealing effect, and debts not being paid back inducing a higher interest rate (see section 5.3).

Entry and exit barriers

In general, there are two strands in the industrial organisation (IO)-literature how to face entry and exit barriers.

First, Bain (1956) defined barriers to entry as 'the advantages of established sellers in an industry over potential entrants sellers, these advantages being reflected in the extent to which established sellers can persistently raise their prices above a competitive level without attracting new firms to enter the industry'. Stated otherwise, entry barriers exist if the most efficient entrant of all potential entrants cannot enter the market.

Second, the Chicago School takes a slightly different perspective. According to Stigler (1968), entry barriers are 'a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry, but is not borne by firms already in the industry. Hence, in this view, the asymmetry of costs between incumbents and potential entrants is important.

The second perspective provides a much shorter list of potential entry barriers than the first one. The most prominent differences between the two strands are the following. In Stigler's view, economies of scale do not form an entry barrier if entrants have access to the same technology. Furthermore, patents are also not an entry barrier in this view as they generate 'economic rents' being understood as opportunity costs. Coherence between the two perspectives is on institutional barriers. Institutional barriers include extra monetary and time costs that a new firm must pay due to policy regulations such as building permits, procedures, environmental policy or food safety approvals. Very broadly, one could distinguish three general types of entry barriers:[8]

- Technological barriers
- Institutional barriers
- Strategic barriers

Barriers with respect to the technological structure of the market are related to economies of scale, for instance the amount of investments required for bringing the costs of operations to a minimum. Strategic barriers are linked to the conduct of existing firms to deter the entrance of a

[8] Note, that those types are closely interrelated and that certain entry barriers can be classified into more than one type. For instance, patents or R&D can be classified as technological barrier but also as strategic barrier.

new firm. As stressed in the new empirical IO literature, those barriers are an important element of strategic conduct of incumbent firms to control competition by limiting entry. Examples are: advertising, predatory pricing and switching costs.

2.3 Productivity, efficiency and welfare

Productivity and welfare

Based on the welfare theorems of economics, markets lead to an efficient outcome (i.e. maximum welfare) as longs as no market failures occur. Welfare can be defined in different ways. Very often, it is defined as the sum of consumer and producer surplus. Welfare can have short term and long-term elements, often expressed in terms of static and dynamic efficiencies.

This document highlights entry, exit and productivity. In general, high productivity (growth) corresponds to high(er) welfare since GDP per capita increases. Higher productivity is not always conducive to more welfare. Higher productivity via more competition does not always lead to more welfare as firms may produce more homogeneous products and less product variety. If consumers prefer the latter, then competition is not conducive to welfare (see Boone, 2000). This section therefore also discusses the relationship between productivity and welfare.

Productivity and efficiency

Productivity is expressed as a ratio of output to input. Since outputs and inputs are typically multidimensional, partial (e.g. accounting for only some inputs) aggregated measures have to be used. Most frequently reported measures of productivity are labour productivity, defined as value added per unit of labour; and total factor productivity, defined as value added per unit of aggregated input, composed of multiple inputs.

Productivity growth encompasses changes in static and dynamic efficiency. The former reflects changes in productivity performance under a given technology. The latter corresponds to productivity improvements due to changes of technology or discovery of new products. The former is, therefore, a short-term concept, while the latter is a long-term concept.

Static efficiency is defined as the condition in which the short term combined welfare of consumers and producers is maximised with production taking place at the lowest cost. Static efficiency is the level at which resources are optimally allocated within the economy (i.e. allocative efficiency) and all firms are on their production possibility frontier (i.e. productive efficiency). The former refers to the allocation of scarce resources among competing firms. Productive efficiency occurs if output is produced in the least expensive way given the current production technology.

However, in practice it is difficult to directly assess static efficiency by measures. Often the level of competition is used as an indirect indicator.[9] If the level of competition is low, (static)

[9] For further discussion see Bennet et al. (2001).

inefficiency of incumbents may be expected. The consequences of fiercer competition can be as follows. Entry of more efficient firms, or even a threat of entry, stimulates incumbents to become more efficient by choosing the best available resource allocation or moving towards the available technological frontier. Both channels reduce static inefficiencies and therefore raise productivity in the industry.

Benett et al. (2001) defines dynamic efficiency as the extent of maximisation of the present value of the consumers' and producers' static utility flow over the long term. Better products and improved production technologies positively affect the expected future revenues, or stated otherwise improve the prospects of higher long-term welfare. Unlike static efficiency, dynamic efficiency is realised as a result of improved technological changes (process innovation), or introduction of radically new, differentiated or better quality products (product innovation). Generally, innovations go hand in hand with productivity growth. We will describe productivity growth in two ways, either by exogenous or by endogenous productivity growth.[10] In the case of exogenous productivity growth, productivity growth takes no resources from the economy, whereas endogenous growth needs resources.

Exogenous growth 'falls like manna from heaven' and may cause entry of a new firm/entrepreneur with a radical innovation or producing improved production techniques based on a brainwave (or 'the hand of God'). Exogenous shocks can also be interpreted as innovations coming from abroad. Given the special characteristics of the Dutch economy, this situation prevails in most cases.

Endogenous growth arises from R&D in process and product innovations. Process innovation decreases production cost (hence lower output prices) and product innovation results in new and improved goods representing higher consumers value. As a result of both types of innovations, long-term welfare increases. All this is valid to the extent that such investments in R&D have indeed resulted in useful innovations and were not wasteful. Hence, innovation and its driving forces play a crucial role for endogenous productivity growth.

Trade-off static and dynamic efficiency

Promoting short-term goals may suppress long-term goals and vice versa. Such trade-off manifests in the form of the trade off between static and dynamic efficiency of production.[11] Theoretical and empirical literature widely point out the existence of a trade off between static and dynamic efficiency meaning the situation where both high static and dynamic efficiency is not attainable (see Aghion et al., 2005a).

[10] As we will explain in chapter 4, the nature of the productivity change is an important explanatory factor behind the mechanism between entry/exit and productivity.

[11] Note that more trade-offs may exist. For instance, new technology increases the demand for higher educated people. This new technology will improve productivity but may also increase the wage inequality between high and low educated people (Skill bias technological change argument). Another trade off might occur between lower entry barriers, more competition, higher productivity but also lower product quality, more pollution or lower work conditions.

As the definitions of static and dynamic efficiency suggest, this trade-off can be captured by the analogy between competition and innovation. On the one hand, more intense competition may stimulate innovation, then a rise in the number of firms may improve static and dynamic efficiency. On the other hand, more intense competition may also reduce innovation. In that case, the loss in welfare in the long run is probably larger than the gain in static efficiency. Additionally, Aghion et al. (2005a) state that there can exist an inverted U-shape relationship between the intensity of competition and innovation. Under a monopoly a low level of innovation can be observed, because not being threatened by potential entrants the monopoly has less incentive to invest and innovate. In the case of the other extreme market situation, when there is cutthroat pricing competition, firms cannot expect economic profits – that is the revenue above variable costs – to invest or to recoup investments in the end. Therefore the level of innovation is also low there. It implies that there is a certain level of competition in between these two extremes, at which dynamic efficiency through innovation is maximised.

Policy makers must keep this potential trade-off in mind when contemplating policies to promote productivity through entrepreneurship and entrepreneurship through firm entry. In case of low (high) static efficiency and high (low) dynamic efficiency, consumer prices are relatively high (low) and investments in new technology are high (low). Here, policy makers should take into account the uncertainty of future expected welfare benefits, costs of reduced static efficiencies and the social preferences for risks and time.

3 Entry and exit: some stylised facts and figures

3.1 Introduction

The aim of this chapter is to present and discuss some stylised facts on firm entry and exit including their costs.[12] We present these stylised facts for the Netherlands and, to put them in an international perspective, for some other countries. Both from a statistical and econometric perspective, we also show results from the relationship between entry, exit and productivity. We summarise the evidence and relate it to economic theory and policy. To the extent that policy may affect entry, its determinants and impact on industry productivity are the issues worth considering by policy makers.

3.2 Entry and exit: empirical evidence

3.2.1 Stylised statistical facts on firm entry and exit

One of the first papers bringing together various facts on entry and exit is Geroski (1995). His study relies heavily on the US experience, though. There have also been several OECD and national studies on entry and exit, some of which we report below in detail. The most basic facts in all those studies are as follows:

- Many firms enter and exit the market each year, at every stage of the business cycle.
- In general, entering and exiting firms have smaller size and lower productivity than average, and a large fraction of new entries do not survive the first few years.
- However, entering firms that survive the first couple of years gain in size and in productivity faster than industry average.
- Market selection seems to be substantially as more than fifty percent of all entrants have already left the market after five years and between 15% and 20% of firms enter and exit the market each year.

Table 3.1 reports basic data on firm entry and exit for a selection of OECD countries collected by EIM, a Dutch consultancy. The average new firm entry rate in the Netherlands for the period was 9.7%, comparing favourably with those in Belgium (7%) and France (6.9%), but lower than in the UK and US. Given the countries observed, the share of new entrants in total employment for the Netherlands is quite small (about 1%), reflecting smaller size of start-ups. The average exit rate in the Netherlands was 6%, with a loss of employment at 0.7% of the total, so that firm entry and exit create a small surplus of jobs. The exit figures for the Netherlands are in line with those reported for Belgium and France but smaller than in Germany, UK and the US Again,

[12] There is a rapidly growing empirical literature in this area as the availability of firm level data has increased. This chapter refers to a selection of it to provide an overview of stylised facts.

comparing the exit rate with the share of exiting firms in total employment, we conclude that exiting firms (in the Netherlands) are smaller than average.

Table 3.1 Basic statistics on firm entry and exit, 1995-2005[a]

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Average
Netherlands												
Entry rate	11.0	10.0	10.0	9.6	10.0	11.0	9.7	8.1	8.0	8.8	9.9	9.7
Share of entries	1.2	1.1	1.1	1.1	1.1	1.3	1.1	1.0	1.0	1.1	1.2	1.1
Exit rate	5.5	5.3	5.5	5.6	5.7	6.1	6.3	6.1	6.5	6.5	6.4	6.0
Share of exits	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.7
Belgium												
Entry rate	8	7.4	7.2	7.5	6.5	6.7	6.5	6.4	6.4	7.2	7.6	7.0
Share of entries	1.6	1.5	1.4	1.5	1.2	1.2	1.2	1.2	1.1	1.3	1.4	1.3
Exit rate	6.2	6.7	6.7	6.6	7.1	6.8	6.6	6.3	6	6.2	6.2	6.5
Share of exits	1.3	1.4	1.3	1.3	1.4	1.3	1.2	1.2	1.1	1.2	1.2	1.3
Germany												
Entry rate	14	12	13	11	11	9.9	9.2	8.9	10	11	9.1	10.8
Share of entries	1.2	1.1	1.2	1.2	1.1	1.0	1.0	1.0	1.1	1.3	1.1	1.1
Exit rate	8.8	8.3	8.3	8.1	8.0	7.3	7.0	7.0	7.1	7.3	7.4	7.7
Share of exits	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.9	0.9	0.8
France												
Entry rate	7.3	6.9	6.8	6.7	6.8	6.9	6.8	6.7	7.8	6.5	6.6	6.9
Share of entries	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.8	1	0.8	1	0.9
Exit rate	n.a.	4.6	7.0	5.8	5.2	5.5	4.3	4.1	4.9	5.3	5	5.2
Share of exits	n.a.	0.6	0.9	0.7	0.6	0.7	0.5	0.5	0.6	0.7	0.8	0.7
UK												
Entry rate	13	13	13	13	13	13	13	13	13	13	12	12.9
Share of entries	2.1	2.1	2.0	2.1	2.1	2.1	2	2	2	2	2	2.1
Exit rate	11	11	9.8	10	11	12	11	10	11	12	9	10.7
Share of exits	1.8	1.7	1.8	1.8	2.1	2.1	1.9	1.9	2.1	2.1	1.6	1.9
US												
Entry rate	10	12	12	10	10	10	9.5	9.5	9.2	9.7	10	10.3
Share of entries	2.1	2.4	2.5	2.1	2.1	2.0	1.9	1.9	1.9	2.0	2.0	2.1
Exit rate	8.7	8.8	9.5	9.0	9.0	8.9	9.2	8.9	8.7	8.8	8.9	9.0
Share of exits	1.8	1.8	1.9	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8

[a] Entry rate is defined as the number of newly created firms or existing firms who started a new business activity divided by the total number of firms in the industry. Exit rate is defined as the ratio of firms who stopped their activities to the total number of firms in the industry. Entries the number of people employed in newly created firms or new activities by existing firms relative to the total number of people employed in the industry.

Source: Own computations based on EIM-database available on their website.

The ‘creative destruction’ theory of new firms driving economic growth implies positive correlation between entry and exit. Indeed, the correlation between entry and exit for Germany, UK and the US turns out to be positive: 0.9, 0.7 and 0.5, respectively. However, for the Netherlands and Belgium it is negative, and for France it is zero. Note that the correlation between entry and exit is negative where the rates are relatively low and positive where the rates are high. That is, lower entry and exit seems to correspond more closely to trend-like or cyclical developments in the economy, whereas higher entry and exit likely allows for more ‘creative destruction’.

Given these basic facts, a lot of research has been done to relate them to industry-specific factors and institutions. Bartelsman et al. (2003) report the results of a comparative study of firm dynamics in OECD countries. The correlation between entry and exit is positive for most countries, in accordance with the ‘creative destruction’ view. Entry and exit rates decline with average firm size in the industry: entry tends to be somewhat higher in the service sector than in manufacturing. Its variation by country is small.

Table 3.2 Entry and exit: comparison US and EU

	Manufacturing		ICT Producing		Non-ICT	
	US	EU	US	EU	US	EU
Exit share of job destruction	24.7	34.3	10.7	24.1	24.9	37.4
Entrant size, % of average	21.0	38.6	6.3	35.7	24.0	40.8
Productivity gap of exiters	10.0	15.4	1.2	9.1	7.9	17.7
Employment share of exiters	18.9	23.1	20.2	31.8	19.8	22.3

Source: Bartelsman et al. (2008).

Table 3.2 illustrates some recent findings on firm dynamics between the US and EU looking at technology grouping (see Bartelsman et al., 2008). Both entering and exiting firms’ characteristics vary between the EU and US. First, entering and exiting firms in the US tend to be smaller relative to the industry average than in the EU. Apparently, it seems to be easier to enter and exit the market in the US compared to the EU. Second, EU firms exiting the market tend to stay longer, in the meantime growing larger in size and lose more in productivity compared with the industry average (see differences in the shares of exiting firms in total job destruction and employment, also productivity gap of exiters). This evidence suggests that it is more difficult to exit the market in Europe than in the US and that more resources are held in exiting firms than in the US. Finally, earlier work of Bartelsman et al. (2003) also showed that new firms in the US grow much faster than in Europe.

New OECD-publication on measuring Entrepreneurship

Recently, the OECD released a new publication (OECD, 2008) with the results of the first round of empirical data collected under the Entrepreneurship Indicators Program (EIP). The creation of the OECD-Eurostat EIP was a joint effort of the OECD, a consortium of 30 industrial democracies, and Eurostat. The EIP establishes multiple measures of internationally comparable data on entrepreneurial activity in 18 OECD countries, based on information produced by national statistical offices according to internationally agreed definitions.

No time-series are available yet. With regard to the Netherlands, the 'employer' enterprise birth and death rates in manufacturing and services are relatively lower than on average in an international perspective. The one-year survival rate in the Netherlands was the lowest among all countries present in the database.

3.2.2 Some evidence on entry and exit costs

Explanations of those differences in firms' characteristics between the EU and US include institutional differences (tighter regulation of product and labour market, more pre-market selection, less money available for start-ups in Europe) and the business cycle. One aspect of institutional environment worth mentioning here is regulation regarding market entry and exit, especially divestment of labour. Countries with lower entry tend to have higher rates of both entry barriers and employment protection.

Availability of figures on regulation issues has been growing over time. Table 3.3 reports four OECD-indices for a selection of countries. The OECD indicator of employment protection legislation (EPL), based on survey responses, includes three components: i) legislative provisions for dismissal, ii) procedural inconveniences involved in dismissal and iii) dismissal notice and severance pay provisions). The lower the EPL the lower the barriers are.

The second indicator is the one for product market regulation (PMR).[13] PMR is the overall (weighted) result of a number of indicators also based on national survey responses on a number of questions including barriers to entrepreneurship. The lower the index, the lower are the barriers. The barriers to entrepreneurship index is a weighted average including regulatory and administrative opacity, administrative burdens on start-ups and barriers to competition, such as legal restrictions.

The EPL-score is between 0.7 (i.e. US) and 3.5 (i.e. Portugal and Turkey) in 2003. Hiding behind the relatively high score for the Netherlands is strict employment protection of permanent workers in combination with a liberal stance towards temporary work. The strict EPL for regular employment in international comparison reflects large procedural inconveniences, a high difficulty of dismissal and high compensation for individuals with long tenures (see Deelen et al., 2006). Germany, however, does not fit this pattern. It has comparably high levels of employment protection and product market regulation with the rest of Continental Europe, and yet its entry and exit rates are comparable to the UK and US

[13] See Conway et al. (2005).

Table 3.3	Some institutional barriers, 1998 and 2003							
	EPL		Product market regulation		Barriers to entrepreneurship		Administrative burdens on start-ups	
	1998	2003	1998	2003	1998	2003	1998	2003
Australia	1.5	1.5	1.3	0.9	1.4	1.1	1.1	1.0
Austria	2.4	2.2	1.8	1.4	1.7	1.6	2.6	2.8
Belgium	2.5	2.5	2.1	1.4	1.9	1.6	1.3	1.7
Canada	1.1	1.1	1.4	1.2	1.0	0.8	1.4	0.9
Czech Republic	1.9	1.9	3.0	1.7	2.0	1.9	2.2	2.3
Denmark	1.8	1.8	1.5	1.1	1.4	1.2	0.5	0.5
Finland	2.2	2.1	2.1	1.3	2.1	1.1	2.0	1.3
France	2.8	2.9	2.5	1.7	2.8	1.6	3.4	1.9
Germany	2.6	2.5	1.9	1.4	2.0	1.6	2.4	1.6
Greece	3.5	2.9	2.8	1.8	2.1	1.6	3.0	2.6
Hungary	1.5	1.7	2.5	2.0	1.6	1.4	2.4	2.3
Ireland	1.2	1.3	1.5	1.1	1.2	0.9	0.9	0.5
Italy	3.1	2.4	2.8	1.9	2.7	1.4	4.6	2.4
Japan	1.9	1.8	1.9	1.3	2.4	1.4	2.1	1.9
Korea	2.0	2.0	2.5	1.5	2.5	1.7	2.2	2.2
Mexico	3.2	3.2	2.4	2.2	2.7	2.2	3.4	3.1
Netherlands	2.3	2.3	1.8	1.4	1.9	1.6	1.8	1.6
New Zealand	0.8	1.3	1.4	1.1	1.2	1.2	1.0	0.8
Norway	2.7	2.6	1.8	1.5	1.5	1.0	1.8	1.0
Poland	1.9	2.1	3.9	2.8	2.8	2.3	3.8	3.7
Portugal	3.7	3.5	2.1	1.6	1.8	1.3	2.1	1.7
Slovak Republic	2.2	1.7		1.4	-	1.2	-	1.9
Spain	3.0	3.1	2.3	1.6	2.3	1.6	3.5	2.8
Sweden	2.6	2.6	1.8	1.2	1.9	1.1	1.1	1.2
Switzerland	1.6	1.6	2.2	1.7	2.3	1.9	2.2	1.7
Turkey	3.4	3.5	3.1	2.3	3.2	2.5	2.7	2.7
United Kingdom	1.0	1.1	1.1	0.9	1.1	0.8	1.0	0.7
United States	0.7	0.7	1.3	1.0	1.5	1.2	0.9	1.0

Source: EPL (version 2) comes from OECD Employment Outlook 2004; Rest of data comes from Conway et al. (2005).

In contrast, the PMR-ranking is relatively better for the Netherlands (10th place). This is also the case for the administrative burdens on startups. Also other European countries like Belgium, France and Germany perform relatively better. However, looking at the result of the barriers to entrepreneurship, the Dutch ranking is relatively worse.

More indicators can be found. For instance, according to Statistics Netherlands (2006 and 2008), the Netherlands ranks relatively high in terms of number of days to open a business in an international perspective. The World Bank produces the ‘Doing business’ index for 181 countries. This index is based on ten indicators including starting a business, getting credit and closing a business. A high ranking on the ease of doing business index means that the regulatory environment is conducive to the operation of business. The Netherlands is in 26th place in the overall ranking, being 10th in closing business and 51st in starting a business.

3.3 Empirical evidence relationship entry/exit and productivity

3.3.1 Statistical evidence of contribution of entry and exit to productivity growth

Industry productivity growth, the change in productivity levels in time, can be decomposed in several ways to calculate the contribution of entry and exit to it. However, no complete agreement in the literature exists as to how to calculate their contribution. Several most popular techniques are described in greater detail in Appendix A.

In general, every method decomposes the aggregate productivity growth of an industry or economy into three main components. The first component is the productivity growth due to the net contribution of entry of new firms and exit of existing firms. The second component measures productivity growth within the existing firms. Finally, the third component calculates the effect on productivity growth due to changes in the existing firms' shares in total employment. It is obvious that the first component is directly related to the effect of entry and exit on productivity growth. But this is also the case for the second and third component. The effect of (threat of) entry and exit affects competition and hence the behaviour of existing firms.[14] Moreover, it also induces reallocation of sources across incumbents.

Putting an estimate to the contribution of entry and exit to productivity growth has been an important empirical question with a range of implications regarding the importance of 'creative destruction' in the shaping of productivity growth. Table 3.4 summarises such estimates produced in a number of studies, along with the contributions of other factors.

[14] One interpretation of this component is simply changes in allocative efficiency due to an arbitrary reason. However, a more tight interpretation, consistent with theoretical models allowing for firm heterogeneity (see chapter 4 for more detail), may also be proposed. While homogeneous firms will retain their market shares whether there is (threat of) entry or not, if firms are heterogeneous in their productivity, changes in their market shares will be an unavoidable consequence of (threat of) entry. Therefore, productivity growth decompositions as we interpret them are consistent with the relevant theory.

Table 3.4	**Empirical estimates for the contribution of firm entry and exit to productivity growth**						
Study	Sample	Period	Dep.var.	Contributions (% total)			
				Entry + exit	Within firm	Reallocation	Covariance
Baily et al. (1992)	US census of manufacturers excl. IT equipment & motor assembly	1972-77	TFP	3.2	30.8	65.9	
		1977-82		4.6	68.7	26.7	
		1982-87		8.0	68.5	23.5	
Griliches and Regev (1995)	Survey of Israeli manufacturing firms	1979-88	Value added	12.0	88.0		
Haltiwanger (1997)	US census of manufacturers	1977-82	TFP	7.8	5.4	22.9	63.9
		1982-87		8.7	43.2	12.6	35.5
van der Wiel (1999)	Business services in the Netherlands	1987-95	Gross output	41.7	8.3	50.0	0.0
OECD (2001)	Business registers, 8 OECD countries	Mid 80s-90s	Value-added TFP	10-40	60-85	0-20	
Balk& Hoogenboom-Spijker (2003)	Manufacturing in the Netherlands	1984-99	Value-added TFP	3.0	87.0	10.0	
Foster et al. (2005)	US census of manufacturers	1982-97	Revenue -based TFP	22.4	29.5	12.3	35.7
			Output-based TFP	22.8	58.6	6.8	11.9
Baldwin & Gu (2006)	Canada Census of manufacturers	1979-89	Gross output	24.0	27.8	48.2	
		1989-99		28.9	31.3	39.8	

We see that the decompositions using constant employment shares in the counterfactual approach (examples include Griliches and Regev 1995; Haltiwanger 1997) tend to ascribe relatively little importance to entry and exit. Their direct contribution is most of the time around 10%; and their indirect contribution, via employment share reallocation, tends to be modest as well. On the other hand, the Baldwin and Gu (2006) decomposition gives greater weight to both the contribution of firm entry and exit and that of output share reallocation. With the total contribution of entry and exit estimated at about 70% (20-30% direct effect and 40-50% market share reallocation), it becomes an important phenomenon for economic research. Instead, holding employment shares constant in the presence of unequal productivity growth between exiting firms and incumbents implicitly allows for some firm turnover; as a result, the direct and indirect contributions of entry and exit become underestimated.

Admittedly, the results in Table 3.4 are not directly comparable due to two issues. First, they are based on different data samples. Second, the contribution of entry and exit also tends to grow with the length of time period, simply because as the time period gets longer, more firms classify as entries of exits. However, Baldwin and Gu (2006) use Haltiwanger and Griliches and Regev's decompositions and the authors arrive at qualitatively the same results. Balk and Hoogenboom-Spijker (2003) also report high sensitivity of their decomposition results to the choice of the method, with smaller contribution ascribed to entry and exit by their preferred Griliches and Regev's decomposition, and greater when an earlier version of Baldwin and Gu's is applied. Thus, the method appears to be responsible for more variance in the results than is the sample.

Having resolved the issue of decomposition method, however, another important driver behind a large variation in the results is differences across industries. Earlier CPB research on the business services sector (van der Wiel, 1999) using Haltiwanger's decomposition finds that within-firm accounts only for a small part of the total productivity growth in that industry, which instead is driven mainly by entry and exit as well as competition among the incumbents. This is in stark contrast with the results by Balk and Hoogenboom-Spijker (2003) obtained using a not so different decomposition method applied over roughly the same period of time, but on manufacturing, not services, data. Baily et al. (1992) and OECD (2001) also report significant heterogeneity in their decomposition results depending on the industry observed. Again, consistently with the results by Bartelsman et al. (2003) on the industry difference in growth rates of new firms, the contribution of entry and exit to industry average productivity growth is stronger in more technologically advanced sectors, such as ICT. It is in those sectors where the importance of entry and exit as a mechanism of spreading innovation is the highest - which is evident from the productivity growth decompositions.

3.3.2 Econometric evidence relationship entry/ exit and productivity growth

Another way to investigate the relationship between entry, exit and productivity empirically is to use econometric techniques. Empirical studies are, however, still scarce due to data availability (e.g. time series), measurement problems and not the least dual causality (see next chapter for more on the causality between entry/exit and productivity).[15] We review three recent empirical studies.

Bartelsman et al. (2004) find a positive and statistically significant correlation between firm turnover rates and productivity growth of incumbents. They also find evidence that higher firm turnover goes hand in hand with higher productivity difference between entrants and firms that

[15] There is a bunch of empirical studies available trying to explain econometrically entrepreneurship (or the number of entrants). See for instance EZ (2007) for an overview. However, our document considers the relationship between entry, exit and productivity.

exit. The authors strongly underline that one should be cautious in drawing causal inferences as these results reflect correlations and not causality. For instance, this correlation may be due to competitive pressure put on the incumbents that induces them to excel in their performance. Or it may be the result of the new entrants-turned-quitters' failure to reach the appropriate productivity level, the reasons for which may be technology- or sector-specific.

Carree and Thurik (2007) investigate the effect of changes in business ownership rate (the number of business owners relative to the total labour force) on three measures of economic performance: employment growth, GDP growth and labour productivity growth. Using country-level data for 21 OECD countries, they find no strong evidence for a negative effect of firm entry on productivity over time, even though entrants do operate below the minimum efficiency scale. A lesson to be learned from this study is that it may be hard to tell between the two contradicting effects of firm entry on aggregate productivity: a reduction due to entrants being less productive than average, and an increase due to competitive pressure put on the incumbents.

An alternative explanation to the lack of correlation between business ownership rate and productivity growth may be that this measure is confounded by factors positively related to productivity but negatively to business ownership. Thus, exploiting economies of scale leads to higher labour productivity and to a reduction in business ownership at the same time. Indeed, Carree and Thurik (2007) find that an "equilibrium" business ownership rate is negatively related to the level of the country's economic development as measured by per-capita GDP.

After correcting for the "equilibrium" business ownership rate (i.e. that predicted by per-capita GDP), Erken et al. (2008) find that the corrected business ownership rate is positively related to aggregate productivity growth in 20 OECD countries. Note, however, that because changes in the business ownership rate are the net result of changes in entry and exit, this study do not differentiate between the individual effects of entry and exit on the productivity performance.

Summing up, the econometric evidence provides an ambiguous correlation between business ownership and productivity growth. Moreover, this correlation does not say anything about causality taking into account that business ownership is considerable persistent over time and do not differentiate between entry and exit. Therefore, chapter 4 takes up this issue of disentangling the relationship between entry/exit and productivity growth from a theoretical perspective.

3.4 Concluding remarks

We have seen that a significant fraction of firms enters and exits the market within a given period of time. These firms are smaller and less productive than average, and there is considerable attrition among the new entries. The surviving new firms, however, grow faster than average, in terms of size as well as productivity. These basic facts notwithstanding, firm entry and exit is a complex phenomenon influenced by industry-specific and institutional factors. For instance, we have seen that a positive correlation between firm entry and exit, predicted by the 'creative destruction' theory, is hardly a universal phenomenon, but instead appears to be influenced by institutions.

In practice, there are two ways to investigate the relationship between entry, exit and productivity. First, productivity growth decompositions help to gauge the importance of entry and exit for sectoral productivity dynamics. With a few exceptions, most of the recent empirical evidence suggests a significant contribution of entry and exit to industry productivity growth. Firm entry and exit may account for between 20% and 30% of the total productivity growth, and 50% may be due to market share reallocation among the continuing firms.

Another way to investigate the relationship between entry, exit and productivity is to use econometric techniques. Empirical studies are, however, still scarce due to data availability (e.g. lack of time series), measurement problems and last but not least dual causality. The econometric evidence provides an ambiguous correlation between business ownership and productivity growth.

With respect to the Netherlands, empirics come up with interesting findings. The Dutch economy features relatively few new firms compared with other countries. Moreover, based on rankings, most regulation indicators like EPL and (components of) PMR suggest that institutional entry barriers are not low in an international perspective.

What can we take away from here with respect to the relationship between entry, exit and productivity? It must be noted that productivity growth decompositions are a simple statistical tool with no underlying economic theory regarding the relationship between entry and productivity growth. In particular, we cannot at this point say much about the effect that the threat of entry might have on productivity growth within the continuing firms, other than observing large differences in the contributions between industries. Similarly, the econometric correlation does not say anything about causality taking into account that business ownership is considerable persistent over time and do not differentiate between entry and exit.

Theories do provide insights in the relationship between entry, exit and productivity. Therefore, we move to the literature review of recent theories to understand the causality in the next chapter. Nevertheless, the lessons of this chapter are as follows:

- Entry and exit rate can be sizeable
- Firms are heterogeneous
- Industry-specific issues and institutions might matter

With respect to the second bullet point, the observed differences in size and performance dynamics between entrants and incumbent firms question the validity of the 'representative firm' approach in neo-classical economic modelling. Not only do researchers need to take account of a cyclical movements in size and productivity among different firms within the same industry, but also provide for different ways in which the performance of a given firm may be improved depending on how long it has been on the market.

4 Theory on relationship entry/exit and productivity

4.1 Introduction

The stylised facts presented in chapter 3 show that firm entry and productivity are connected. Yet, as discussed, these methods do not explain the observed relationships between entry/exit and productivity and do not say much about causality between these phenomena. Understanding of the underlying mechanisms is crucial when designing policies to stimulate productivity growth. Therefore, in this chapter we review main theoretical papers that provide insight into these matters. Our aim is twofold: to give an overview of the literature in a systematic manner, and to see how policymakers can use this literature.

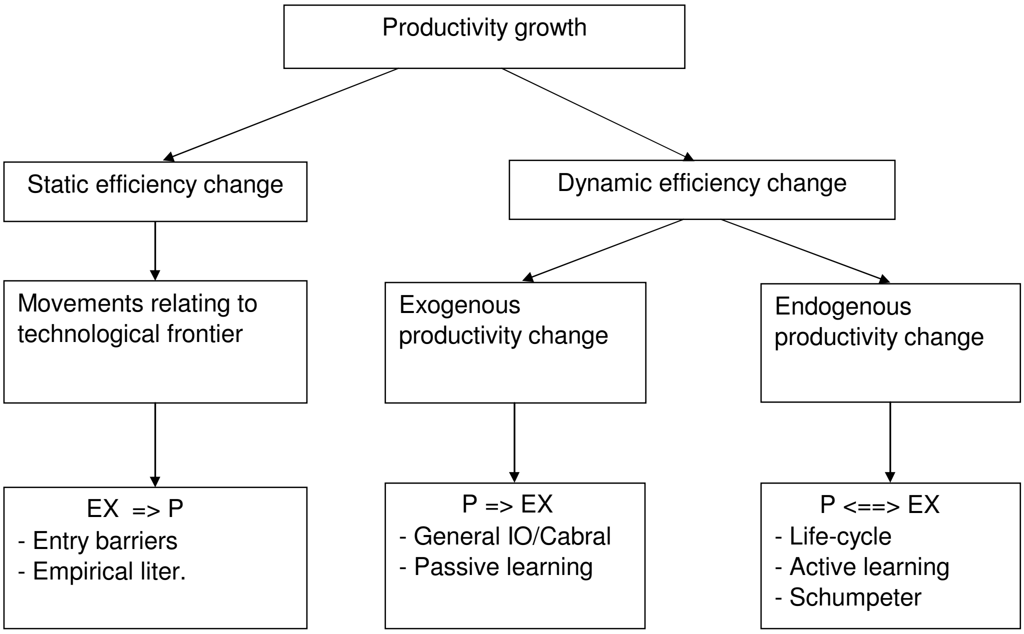
With respect to the first aim, our literature review builds upon the survey presented in EC (2005). This paper discusses effects of entry/exit on industrial productivity performance, as well as the determinants of both. However, it concentrates on the empirical side of the issue, giving only a brief overview of theoretical models that provide the basis for empirical analysis. In particular, EC (2005) focuses on two effects: (i) the effect of market reforms on entry and exit, and (ii) the effect of entry and exit on macro-economic outcomes, such as productivity and innovation. Our document complements that study by incorporating insights from the theoretical literature on the *interaction* between entry/exit and productivity. In addition to four types of theoretical models mentioned in EC (2005, pp. 11-13), we also review a few other models, highlighting main assumptions of each model.

As we will see, different theoretical assumptions may yield different mechanisms of interaction between entry/exit and productivity, and hence, different implications for policy. By highlighting main similarities and differences between modelling approaches, we hope to better understand the scope of their applicability to the case of the Netherlands. In this way, we achieve the second aim of our literature survey by providing the answers to the following four questions related to policy making:

- What exogenous factors affect entry/exit of firms and the relationship between entry/exit and productivity?
- What is the role of entry and exit in determining productivity growth of an industry?
- Does the theory address productivity at the level of a firm, an industry or the macro-economy?
- How does it account for costs associated with firm entry and exit, such as adjustment costs?

Figure 4.1 shows the structure upon which we build the review of the theoretical literature. Productivity growth entails changes in static and dynamic efficiency (see chapter 2 for definitions). Since dynamic changes can be modelled as exogenous and endogenous changes, the ‘tree’ in figure 4.1 splits into three main branches. We will show that these branches correspond to three possible causal relationships between entry/exit and productivity (shown in the bottom row in figure 4.1).

Figure 4.1 **Systemization of literature on entry/exit (EX) and productivity (P): main channels of productivity growth**



Note that our analysis focuses on the interaction between entry/exit and productivity. Therefore, the literature review presented in this chapter focuses on this important relationship, omitting or only briefly commenting on other potentially relevant but less directly related issues. In particular, we do not review the broad literature on innovation (R&D, patent races, pre-market selection, effect of market structure on incentives for innovation, etc.), the IO literature that focuses on entry incentives (entry and market structure, entry and ex-post competition, entry and the ability to appropriate investments, etc.), or other indirectly related literature.

In the remainder of this chapter, we review the literature in each branch of figure 4.1 (in sections 4.2 to 4.4) in more detail with a special focus on their policy implications. At the end, in section 4.5 we summarize the main insights from this literature.

4.2 Static efficiency change

In this section we look at the literature on the relation between entry/exit and static efficiency. The static approach to efficiency considers technology as given. Under this approach, the market is called efficient if consumer surplus reaches its maximum level, and firms gain normal profits. Otherwise, the market is characterised by inefficiencies.

Two main forms of static inefficiency are allocative and productive inefficiency. Production efficiency characterises the firm production level relative to the production possibility frontier. In other words, a firm is productively inefficient if it produces less than it is possible to produce, given the technology level available. Allocative inefficiency arises because of wrong allocation of inputs and products within the economy. Reducing both types of inefficiencies typically increases aggregate productivity.

Models

Two assumptions traditionally used in theoretical models are: perfect information and profit maximising behaviour of firms. Under these assumptions, firms never waste inputs or have a slack in production. Taking input prices and the output level as given, they choose an optimal input mix to achieve the maximum productivity for this output level. Hence, this literature neglects the possibility of productive inefficiency, and rather focuses on allocative efficiency. More modern literature, especially the IO literature that deals with regulatory issues (see, e.g. Laffont and Tirole, 1993), also incorporates informational constraints and agency issues. For example, information asymmetry leads to moral hazard and adverse selection problems that affect incentives of the firms in a market economy. Such models do allow for productive inefficiency.[16]

Allocative inefficiency arises in models with market imperfections (e.g. monopoly or oligopoly markets) or constraints (e.g., physical, financial or regulatory constraints), because of which the firm produces below its efficient output level or cannot buy efficient amounts of inputs at efficient prices. For example, a firm with market power can restricts supply in order to generate a higher profit. In case of scale economies in production, operating at a suboptimal scale results in lower productivity. Also for the input market, market imperfections may distort the price level of inputs in the production process, leading to suboptimal allocation of inputs among firms, unnecessarily raising production costs and thus decreasing both the productivity of individual firms and total productivity in the industry.

Results and exogenous factors

Competition contributes to static efficiency improvement to the extent to which it helps to reduce imperfections and align incentives within firms. In particular, entry of new (more efficient) firms increases static efficiency and hence productivity via two channels. First, an efficient entrant will gain market share, causing a change in the allocation of production, resources and in market prices in an industry. Second, new entrants, and even the threat of entry, will create competitive pressure on inefficient incumbents, stimulating them to catch up with the production frontier or forcing them out of the market. In Figure 4.1, this effect is

[16] Actors insulated from risk may behave differently from the way they would behave if they were fully exposed to risk. Moral hazard arises because they do not bear the full consequences of their actions, behaving less carefully than they otherwise would.

shown as EX⇒P, meaning that in this branch of the tree the causality runs from efficient entry (or its threat) to productivity increase at the firm level as well as at the industry level.
Note however, that in a competitive environment, entry will eventually cease, since only the efficient number of firms will survive in equilibrium. When the number of firms becomes too large, 'diseconomies' of scale, scope, or density arise, which reduces productivity. Competition, therefore, forces some of the firms to exit. The equilibrium number of firms depends on characteristics of the demand function and on the production technology.

Empirical verification

The empirical literature shows that a large gap can occur between the observed productivity and the productivity that would be consistent with profit optimisation. This gap is the so-called X-inefficiency, a term introduced by Leibenstein (1966). Numerous empirical studies provide evidence of large discrepancies in production efficiency of firms operating within the same industry, e.g., Barla and Perelman (1989) on the efficiency of airlines, or Jamsab and Pollitt (2001) on the efficiency of electricity companies. Most of these papers use frontier methods, such as data envelopment analysis and stochastic frontier analysis to measure efficiency differences across firms. These papers often focus on the measurement of efficiency levels or changes rather than on their explanation. However, when the data cover the period in which the industry was undergoing institutional reforms (e.g. deregulation and introduction of competition in the public utility sectors), the efficiency changes occurring in this period can be interpreted as driven by these reforms. In this way, this literature provides us with new empirical insights on the effects of competition on static efficiency. This includes, for example, the effect of abolishing legal monopolies and allowing entry in some industries, such as telecom, or the effect of liberalisation of the energy industry. Typically, these papers show that allowing for more competition improves efficiency of incumbents, which is in line with economic intuition as well as with expectations on the reforms. See for example, Steiner (2001), Bushnel et al. (2005), Fabrizio et al. (2006) on efficiency improvements in the electricity industry after reforms, probably driven by changes in incentives for firms in a competitive environment (which includes a competitive pressure from potential entrants).

While most papers find positive effects of reforms in general, the disentangled effect of entry does not always point in the same direction. For example, Driessen et al. (2006), who analyse the effect of reforms on efficiency of European railways, find a positive effect of competition in general, yet a negative effect of free entry on efficiency. Thus, although competitive tendering improves productive efficiency, free entry lowers productive efficiency in the railways sector.[17] A possible explanation for the latter being that free entry may disable railway operators to reap economies of density in railways.

[17] This result might be related to the issue of competition on versus for the market. We leave this question for future research.

Conclusions

With respect to the four questions put forward in section 4.1, we conclude the following. The 'static' branch of the productivity literature considers technology as given. Improvements in static efficiency are driven by better utilisation of existing technologies, not by technological changes. Static efficiency improves because of relieving certain constraints (e.g. changing institutional settings, allowing entry, etc.) which changes the allocation of production and creates competitive pressure, improving incentives for firms. A more competitive market structure typically leads to a better productivity performance. Therefore, if entry is efficient and increases competition it should also serve as a mechanism to increase productivity. In the empirical literature, the overall effect of competition on static efficiency is often found to be positive. However, there is no such a general result on the pure effect of entry/exit on efficiency, probably because the room for entry depends on the technology and the market size. In this literature, the measurement of efficiency is most often done at the firm or industry level, not at the aggregate economy level. This literature typically neglects adjustment costs.

4.3 Exogenous productivity growth

Exogenous productivity growth occurs when firms do not spend their own resources to get access to innovations, yet firms' productivity increases. In this case innovations may arrive from abroad or from other industries providing complementary products that are important for the industry in question. Under an exogenous productivity shock, firm's individual productivity is a determinant of its entry or exit decision: P⇒EX. In the following sections we discuss two types of models: a fundamental IO-model for homogenous firms based on Cabral (2002) and the passive learning models that assume firm heterogeneity based on Jovanovic (1982) and Hopenhayn (1992).

4.3.1 General static industrial organization models

Model

A standard model in the IO-literature relates the number of firms to the productivity level in a given industry which produces a homogenous good. Cabral (2002) observes that in such markets, a limited number of firms supply that good. The aim of the model is to explain this observation assuming that firms are homogenous with identical productivity levels.[18] The model shows that the condition for a limited number of firms in a market of a homogenous good is production under increasing returns, in this model denoted as fixed costs. These fixed costs and the marginal costs are the exogenous factors which measure the state of the technology. There are also demand-related exogenous factors for firms, namely the price elasticity of demand and

[18] More details of the functioning of this model can be found in a memo that is available on request.

market size.[19] In equilibrium, the endogenous variables, such as the number of firms and the productivity level of the industry depend on the interplay between all exogenous factors. Given these ingredients the relation between the number of firms and productivity can be explained as follows.

A lower price leads to larger demand for the products of the industry. Firms maximise their profits given the production of their competitors and they enter as long as they can make nonnegative profits. Total costs of each firm are equal to the marginal costs times the production volume plus fixed costs. The model is applicable to our research question because (a part of) fixed costs could be attributed as entry costs and all firms are considered as entrants. The equilibrium number of firms is determined as market size divided by the production volume per firm. In the equilibrium, there is no entry. Then, a technological improvement becomes freely available for all firms and a new equilibrium is reached, ignoring the costs of adjustment. It is possible then to derive the effects of an exogenous productivity change on the number of firms and productivity using comparative static analysis. As the model is static, it does not show entry and exit of firms during the process of adjustment from the number of firms in the original equilibrium and the number of firms in the new equilibrium. By definition, however, the change in the number of firms between both equilibriums can only be achieved with entry and exit.

Results and exogenous factors

The main exogenous factor in Cabral's model is technological progress which manifests itself in two ways: by lowering marginal costs, or by lowering fixed costs. Both types of improvement can be interpreted as process innovations, because the same products are made more efficiently. The greater the decline in the costs, the more radical the process innovation is. The model assumes homogeneous goods, hence product innovations are absent. We do not discuss the impact of changes in the demand-related exogenous factors on the number of firms and the productivity of the industry as this section deals with changes in technology or productivity.[20]

The model's findings with respect to the impact of technological change on the number of firms and productivity depend on whether technological changes affect the marginal costs or the fixed costs. Table 4.1 summarizes the impacts of exogenous (productivity) change on productivity of the industry. Consider a decline in the marginal costs (for example, due to energy saving technology) first. Due to perfect competition, none of the firms is able to influence the market price, thus the price of output drops to the new, lower level of marginal costs, increasing demand. Since output per firm depends on economies of scale, a decrease in the marginal costs does not affect the size of existing firms. Therefore, new firms entering the

[19] Market size can be measured by population or maximum consumer expenditures.

[20] It should be stressed however that the levels of the demand-related exogenous factors are important for the size of the impact of technological improvements on the number of firms and the productivity level of the industry, because of the interplay between all exogenous factors.

market satisfy all the extra demand. Due to constant returns to scale, the productivity of the industry increases by exactly the same amount as the fall in the marginal costs.

Technological progress is equally relevant to fixed costs as to marginal costs. Consider as an example the miniaturization of fixed assets due to the application of information technology or electrical production technologies in the steel industry. A drop in the fixed costs reduces total average costs, causing the price of output to fall and demand to rise. As in the earlier case of a drop in the marginal costs, greater demand invites new firms to enter the market.

A drop in the fixed costs has an additional effect on firm entry compared to the case with lower marginal costs. Unlike the marginal costs, a decline in the fixed costs causes the optimal firm size to fall. This happens because lower fixed costs now can be spread over the same amount of output, which, assuming the marginal costs stay the same, shifts the minimum point on the average costs curve to the left. So, not only is there firm entry due to greater demand, but also due to the lower optimal firm size, the entrant is now smaller and more entrants will appear on the market: hence the extra + in the respective cell of Table 4.1.

The implications of the two kinds of technological shocks for aggregate productivity are straightforward. Both a drop in fixed and marginal costs reduces the total average costs proportionally to their shares in the total. The benefits from either reduction are fully passed on to the customers because firms enter as long as they can make profits. Because of the market assumed to be perfectly competitive, a given fall in the average costs translates completely in a lower price. In other words, industry productivity grows by the same amount as the fall in the average costs.

In sum, entry and exit do not affect aggregate productivity in Cabral's model. Instead, an exogenous change in production technology causes both entry and exit and aggregate productivity to change. However, firm entry or exit remain two important mechanisms through which the industry moves to a new competitive equilibrium and productivity gains are achieved in full.

Table 4.1 Cabral case: impacts of changes in exogenous factors

	Number of firms	Productivity industry
Lower marginal costs	+	+
Lower fixed costs	++	+

Empirical verification

Cabral's model explains the stylised fact that in markets of homogenous goods there is a positive relation between technological change and the increase in the number of firms (and consequently entry and exit in the period of adjustment). It also explains the increase in the productivity level of an industry between the equilibrium before the technological improvement and the equilibrium after the improvement. However, the model lacks analytical depth in predicting how many firms enter and exit the market after a exogenous change in technology. For example, an improbable scenario under which all incumbents are replaced by more efficient entrants is still consistent with the model. Therefore, being generally in line with the empirics, Cabral's model leaves room for other (exogenous) factors of productivity growth and the adjustment costs, which it ignores. Moreover, the model assumes that firms on the market have identical productivity levels being inconsistent with empirical evidence of large firm heterogeneity.

Conclusions

In Cabral's model the main exogenous factor is a productivity shock that lowers the marginal and/or fixed costs. The endogenous variables are the number of firms and the productivity level of the industry due to this technology change. A change in marginal costs does not influence optimal firm size, while after a drop in fixed costs firm size is smaller in the new equilibrium than in the original equilibrium. The number of firms, and thus net entry, is endogenous and as such has no impact on other variables. Instead, firm entry is a mechanism through which the benefits of technological growth arise.

The model learns that in markets of homogenous goods, fixed costs (or more generally production processes with increasing returns to scale) are a condition to determine the number of firms and the productivity level of the industry. In markets with heterogeneous products, fixed costs are no condition.

4.3.2 Passive learning models

The theoretical predictions in Cabral's model are based on a heroic assumption that better technology is available instantly and to all firms in a market. We consider passive learning models to see how firm heterogeneity in terms of productivity may be relevant for theoretical predictions regarding the link between firm entry/exit and industry productivity in the presence of exogenous technological changes. The notion 'passive learning' refers to the process when firms entering the market learn their true productivity relative to their competitors' and decide whether to continue or not. Therefore these models are applicable for modelling firms' entry/exit decisions on the basis of their productivity relative to competitors'. An extra advantage of such models over Cabral's is that they can predict not only how many more firms will be on the market but also which firm will stay, and which will quit.

Models

Two prominent papers, Jovanovic (1982) and Hopenhayn (1992), build up theoretical models to establish background for further research taking firm heterogeneity into account, and in particular to explain the often observed variation of the growth rates of firms with their size. They argue that firms differ in size not only because of the initial capital but also because, with time, some of them learn that they are more efficient than others and adjust their size accordingly. This learning process takes place as firms face exogenous individual productivity changes upon which they evaluate whether to stay in the industry or leave it. Furthermore, these models take into account other exogenous changes in industry characteristics, which may also affect entry and exit.

Passive learning models assume that (i) real productivity is gradually learned, (ii) firms with negative expected discounted profit exit, and (iii) the only investment is a one-time entry cost. The firms are ex ante homogenous and operate on the market for a homogenous good. However, heterogeneity in efficiency is introduced as each individual firm learns its true productivity after entry. This heterogeneity determines firms' decisions on whether to continue or not. Firms are price-takers in the input and output markets, and there is only one input - labour. The markets are perfectly competitive. Both Jovanovic's and Hopenhayn's models are able to explain the above-mentioned stylised facts, but Hopenhayn's model is better able to distinguish between the effects of different factors on entry and exit. In particular, he focuses on the cost of entry, fixed costs, input demand, and the characteristics of the stochastic process that individual firms' productivity shocks follow. The general equilibrium model of Hopenhayn and Rogerson (1993) also considers adjustment costs.

Results and exogenous factors

The main findings of passive learning models are in line with some stylised facts highlighted in chapter 3. Basically, continuing firms are larger and more efficient than those that exit, implying a positive contribution of entry/exit to industry productivity growth. Table 4.2 reports the effects of a selection of exogenous factors on entry and exit and other important variables. Hereafter, we discuss how entry and exit interacts with industry productivity. For policy implications, we are particularly interested in entry costs and adjustment costs.

Table 4.2 Passive learning models: impact of changes in exogenous factors

	Minimum productivity to survive	Entry/exit rate	Firm size	Firm productivity	Productivity of industry
Lower entry cost	+	+	Large firm - small firm -/+ average +/-	+	+
Lower adjustment costs	+	+	?	+	+
Higher demand (under fixed input price)	+	+	+	+	+
Higher distribution of entrants' productivity shocks	+	+	?	+	+

Let us consider the *costs of entry* first. Given a realised distribution of productivity levels across firms, a firm will enter the market if its expected profit net of the costs of entry is nonnegative. Therefore, a lower entry cost implies that a firm with a lower perceived productivity level can enter the market, hence more firms will choose to enter. However, with more entry happens more exit, as growing competition decreases price which drives up the minimum level of productivity required to stay on the market. Thus, even though it is their individual productivity that causes firms to enter or exit the market, lower entry costs ensure that tougher competitive selection through free entry increases industry productivity by pushing inefficient incumbents off the market.

Adjustment costs are considered in Hopenhayn and Rogerson (1993) as a form of 'tax on job destruction' applied to individual firms. Higher adjustment costs lead to less job creation, because firm expect that later on there will be less need for job destruction. In our understanding, less job destruction can reflect to less entry and less job destruction to fewer exits. Therefore we may conclude, that adjustment costs reduce firm exit and therefore, holding the market size constant, entry, thus diminishing their positive contribution to productivity level as described above. They have a further negative effect on industry productivity because of the distortion they cause to the allocation of resources (in this case, labour) across firms. Thus, in the presence of adjustment costs, firms that are more productive than expected will acquire less labour than optimal, so that their weight in total output will be too low. Equally, the less productive firms will shed less labour than optimal because it is costly for them to do so, thus their weight in the total will be too high. This inefficient use of resources leads to underproduction and so higher prices. Therefore consumers get worse off so total static welfare is not optimal.

The remaining two factors in Table 4.2, higher demand and distribution of entrants' productivity shocks, behave in an expected manner. Higher demand, at a given input price, causes existing firms to increase production volume, but at the same time induces more entry thus raising the minimum productivity level required to survive. A shift in the distribution of entrants' productivity changes to the right should increase their size (although there is no explicit result in the papers) and make them more competitive compared to the incumbents, thus raising the minimum productivity level. More firms will enter the market once the productivity level they expect increases.

Empirical verification

Empirical studies based on passive learning models support some prominent stylised facts on entry and exit.[21] Similarly to the theoretical models, Evans (1987) and Dunne et al. (1989) find that firms' growth, the variance of firms' growth and the probability of exit decrease with firm age. Disney et al. (2003) examine the contribution of 'internal' restructuring (such as new technology and organisational changes among survivors) and 'external' restructuring (exit, entry and market share changes) to productivity growth. They find that 'external restructuring' accounts for 50% of establishment labour productivity growth and 80–90% of establishment TFP growth, and that competition is an important determinant of internal restructuring.

[21] For a summary of empirical literature see Caves (1998) and Bartelsman and Doms (2000). For empirical literature on developing countries see Tybout (2000).

Conclusions

The passive learning models aim to explain firm dynamics within an industry and how entry and exit affect industry efficiency. These models assume that firms receive exogenous productivity changes. These changes determine expected profits and trigger entry and exit. Models find that entry of more productive firms and exit of less productive ones as well as the productivity growth of incumbents contribute to higher industry productivity. As Table 4.2 shows, lower entry costs and lower adjustment costs lead to higher entry and exit and higher productivity for firms.

4.4 Endogenous productivity growth

In endogenous growth models firms spend their own resources as a response on market opportunities to come up with technological progress. More precisely, the accumulation of knowledge is the underlying source of sustained growth supported by large and persistent spillovers to other agents of the economy. Human capital (including education, on-the-job-training and learning by doing), scientific research, process innovation and product innovation contribute to more accumulated knowledge reflected in technological progress. This document focuses on innovation in connection with entry and exit. Innovations affect productivity and also the entry/exit decisions of firms, and both determine innovations in the coming periods. Therefore there is interdependence between productivity and entry and exit (P⇔EX). In this section we first review two IO-models that focus on firm and industry dynamics: life-cycle models (section 4.4.1) and active learning models (section 4.4.2). Then we discuss two Schumpeterian growth models that focus on aggregate productivity (section 4.4.3).

More endogenous growth models, but neglect overall impact of entry and exit

There are more endogenous growth models available, mostly referred to as AK-models and Romer's product variety models. However, unlike those growth models, Aghion and Howitt (2006) argue that Schumpeterian growth paradigm produces testable predictions as to how competition (including impact of entry and exit) should affect growth. Moreover, those models ignore the importance of taking account of the country's or sector's distance to the technological frontier. With respect to entry/exit and growth, the AK-models do not say anything on how growth is affected by competition and entry policy as up to now those models assume perfect competition.

Exit is always bad in the product variety models as it reduces the economy's GDP by reducing the number of varieties, whereas it has a positive effect on innovation and productivity growth in Schumpeterian theory for incumbents but also for the total economy. Entry is always growth enhancing as it increases product variety. Besides the idea of distance to the frontier, Product variety models 'ignore' the escape competition effect. Simply because in those models, entrants innovate whereas the escape competition effect requires that incumbents perform innovations.

4.4.1 Life-cycle models

The life-cycle theory explores the idea that industries follow a development cycle, also called product life cycle, or 'PLC'. This line of research goes back to 70ths. Two influential papers in this area belong to Abernathy and Utterback (1975 and 1978), who put forward the argument that the PLC is driven by the way in which new technologies evolve. Gort and Klepper (1982), distinguished five phases of the industrial development and provided empirical evidence on them. Another important paper in this area belongs to Geroski (1995), who also argues that a life-cycle approach is consistent with many stylised facts about entry and exit. Klepper (1996) provides a formal model explaining the PLC phenomenon, which we discuss below in more detail.

The model

The model incorporates two types of innovations: product innovations and process innovations. With respect to product innovations, it is assumed that each firm has some 'expertise'. This expertise and the firm's investment in product innovation determine the probability with which this firm will discover a new version of the product in each period. The firm brings this new version of the product to the market, where extra demand for this version arises. Therefore, the overall demand for the product shifts upwards. Since only one firm produces the new version of the product initially, this firm extracts monopoly profits from the extra demand during one period. However, the incumbents constantly monitor the achievements of their rivals, therefore, in the next period all incumbents are able to produce this version of the product, and the previous invention becomes a part of the standard product so that the monopoly rents disappear.

In addition to investing in product innovations, firms can also invest in process innovations and in capacity expansion. Process innovations allow firms to reduce the (constant) marginal cost of production. Capacity expansion allows them to increase production.

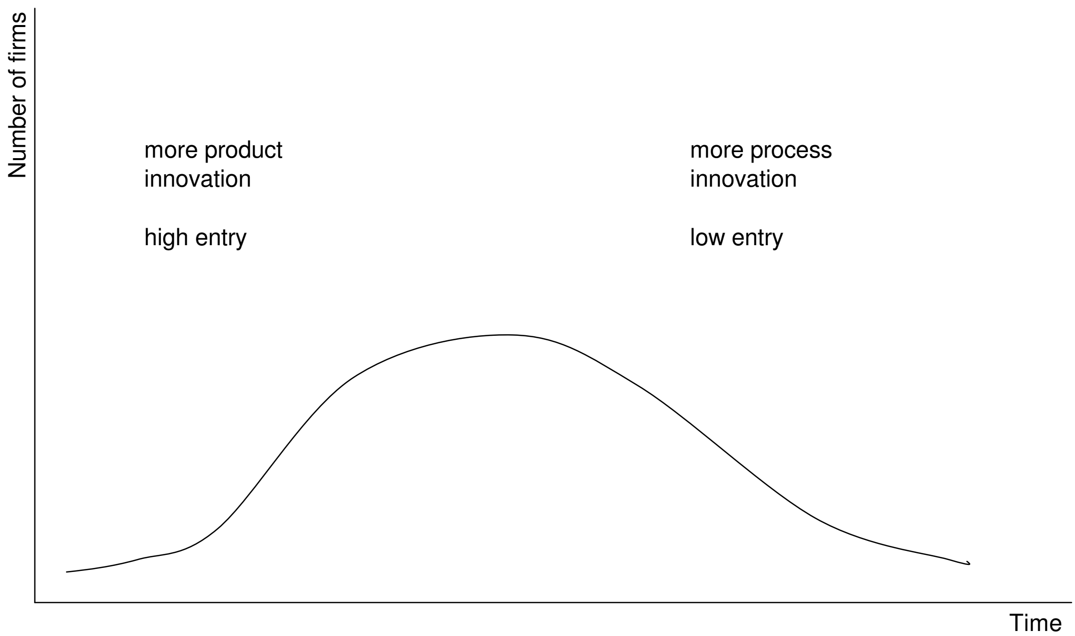
The total costs of the firm, therefore, include: production costs, investment costs in product and process innovations, and the 'adjustment costs' of capacity expansion. All these costs affect production decisions and are endogenously determined, implying that both productivity and entry/exit decisions are endogenous in the model. In particular, entry is driven by successful product innovation (which increases demand of the industry), and exit is driven by unsuccessful product innovation. The adjustment cost is assumed to be convex (as we will see later, this assumption strengthens the advantage of early entrants).

Results and exogenous factors

The model considers firm dynamics within one industry and focuses on differences between 'cohorts' of firms. Since the expected return on innovation is the same for every entrant, all the entrants to the market spend the same proportion of their budgets on *product innovation*. The firms of the same cohort (who entered at the same moment k and still stay in the market in period t) have the same level of production and *product* and *process innovation*, and the same incremental profit from the basic product. Productivity is inversely related to the unit cost of production of the standard product, which depends on the level of investment in process innovation.

The model generates several important results with respect to firm entry. First, entrants introduce relatively more product innovations, while established firms engage in relatively more process innovations (see Figure 4.2). This result arises because in the model a new firm enters the market only if it discovers a new version of the product. Hence, its investment mostly goes mostly into product innovation rather than process innovation. In contrast, incumbent firms are already on the market producing existing versions of the product. As well as creating new product varieties, they also earn from producing the basic product cheaper and save from spreading a fixed amount of the process innovation costs on the growing production volume. Therefore, incumbents spend larger shares of their total investment on process innovations than entrants do.

Figure 4.2 Industrial life-cycle



Second, the model reflects the advantage of early entry. The presence of convex adjustment costs excludes the possibility of an instant growth, implying that firms that entered earlier will have grown larger at any given time. Established firms are therefore larger than new firms and, reinforcing the previous result, have a larger incentive to engage in *process innovations*, which are driving aggregate productivity increases in the production of the standard version of the product. They gain more from process innovation as the decrease of the marginal cost of production can be spread over larger volumes of output. Hence, differently from other theories considered in this chapter (e.g. see section 4.3.1), this theory generates a positive relation between the firm size and productivity.

Since earlier entrants have an advantage, it becomes more and more difficult for new entrants to expand in this market. Therefore, as the surviving firms of earlier cohorts (those who survive) become larger, the threshold value of the innovative ability for successful entry increases. So, apart from institutional barriers to entry, which can be controlled through policy interventions, entry costs and the very incentive to enter may also depend on the stage of the industry's life cycle. Both the number of entrants and the total number of firms in the industry are always decreasing at the end of the cycle. Therefore, the firms who entered in the very beginning of the cycle and had the highest expertise are always larger than the subsequent entrants and earn greater profits. Hence, the model predicts that 'these firms will increase their market shares over time and consistently earn supra-normal profits' (Klepper, 1998).

Since the amount of investment in product and process innovation changes over the life cycle, productivity of each firm also changes over the cycle. Although the model does not focus on implications for individual productivity, its intuition helps us to construct a likely pattern of firm productivity change over time. Thus, in the beginning, there is a strong productivity growth on the back of product innovation. As this source of productivity growth gets 'exhausted', productivity growth abates, showing only incremental increases due to process innovations. As for the aggregate productivity growth in the industry, it also is likely to decrease over the life cycle, because i) the incumbents invest more in process, rather than product, innovation, and ii) with fewer firms on the market competition becomes less fierce.

Empirical verification

The theoretical model in Klepper (1996) is consistent with the empirical evidence on the PLC. As said, this paper preceded by the literature that investigated industrial dynamics in different manufacturing industries and formulated the idea of the PLC. These earlier papers documented stylised facts, which Klepper (1996) aimed to replicate in his model in order to provide a theoretical justification for the PLC theory.[22]

[22] In particular, Klepper (1996) focuses on the following PLC facts: 1) At the beginning of the industry the number of entrants may rise over time or it may attain a peak at the start of the industry and then decline, but in both cases the number of entrants eventually becomes small. 2) The number of producers growth initially and then reaches a peak, after which it declines steadily despite the growth of the industrial output. 3) Eventually the rate of change of the market shares of the largest firms declines and the leadership of the industry stabilises. 4) The diversity of competing versions of the product and the number of major product innovations tend to reach a peak during the growth in the number of producers and then fall over time. 5) Over time, producers devote increasing effort to process relative to product innovation. 6) During the period of growth in the number of producers, the most recent entrants account for a disproportionate share of product innovations.

Conclusions

The model endogenises both firms' entry/exit decisions and productivity, as well as the level of product and process innovations. With respect to exogenous factors, a higher level of 'innovation expertise' increases the probability of the firm's success in product innovations. In addition, some technology and demand-related parameters play a role in determining the resulting productivity level in this model. The model also incorporates adjustment costs.

This model focuses on industry dynamics and firms' investment in innovations, not on aggregate productivity in the industry or in the economy as such. Therefore, the model explains industrial dynamics itself (entry-exit decisions and the evolution of the number of firms, their market shares and investment in innovations over the life-cycle), rather than derives implications with respect to productivity growth. However, following the intuition of the model, it is most likely that initially entry goes hand in hand with higher productivity, whereas at the end of the industrial life-cycle the number of entrants is decreasing while productivity gradually increases to its limits. Another drawback is that it disregards firm exit costs, which may be in practice substantial. Therefore, the main policy implications that we can derive based on this theory concern the relation between policy and industrial dynamics, rather than productivity. When looking at firm entry in a particular industry, policy makers should take into account the stage of the industry's life cycle. There is a naturally high level of entry at the early stages, and low at the end. Therefore, entry should not be restricted at the beginning of the cycle, so that product innovation may take off freely. Also, there is no point in encouraging entry at the end of the cycle, when it is naturally low.

4.4.2 Active learning models

Active learning models aim at explaining stylised facts based on firm-level data and provide theoretical support to test empirically (i) how simultaneous entry/exit may occur, (ii) why similar market characteristics (such as the efficiency of firms or level of investment) or similar exogenous events may lead to different market outcomes, and (iii) how the industry structure may change constantly. The term 'active learning' refers to a process where firms change their pre-entry strategies and investment decisions in response to the feedback from the market. The process of active learning differs from passive learning in that in the former firms spend their own resources on innovation and therefore are able to actively determine their productivity path, whereas in the latter firms only learn but do not influence their real productivity over time.

Models

The most significant paper about active learning is Ericson and Pakes (1995).[23] Their paper is based on a stochastic model of entry and growth of a firm that actively explores its economic

[23] A companion paper, Pakes and McGuire (1994) extends the analysis to heterogeneous product markets. Both papers seek for a very general dynamic (Markov-perfect) Nash equilibrium. To interpret the general findings the authors run simulations for arbitrarily chosen parameter values. See these results later.

environment. The firm invests in order to increase its profit in a competitive environment within and outside the industry. The stochastic nature of the model arises from individual productivity shocks carried out by firms' investments, creating heterogeneity among the firms. Investment can relate to process and product innovation, and both determine firm productivity. It also creates firm-specific uncertainty: the value of the firm's current state is observable, but its changes due to investment follow a stochastic process. The firm's profitability depends on its own output and price as well as on the output and prices of its competitors. Therefore, investment may be a means for the firm to change its own quantity or price or to change its (potential) competitors' strategic variables to create entry threat. Similarly to the passive learning models, entry and exit are determined endogenously. Entry (exit) takes place when the discounted expected profit is larger (smaller) than the opportunity costs of being in the industry.

Results and exogenous factors

Ericson and Pakes's (1995) general model finds that (i) entry is finite in any period and reduces as competition gets stronger, (ii) 'all firms die', meaning the lifetime is finite, but (limited) growth is possible, and (iii) the industry structure is not constant over time, though there is a certain stochastic regularity in it. They also find that changes in the market structure do not depend on the initial market situation.

To illustrate the complex and rather technical general model, Ericson and Pakes (1995) and Pakes and McGuire (1994) run simulations for arbitrarily chosen parameter values. They assume industries with a size of 5 or 6 potential firms at each period, however these firms do not necessarily have to be the same over time. They find that under similar circumstances, when products are homogenous there are generally 1 or 2 firms active in the market, while 3 or 4 firms produce differentiated products simultaneously. The difference in the number of firms is due to the existence of potential market niches created by product varieties. Under product differentiation, a firm is able to monopolize the market of its own product version, therefore (ex post) competition does not drive down prices as intensively as in the case of homogenous goods.

Both papers find that when entry costs are low, entry rate is high, but many firms survive only a few periods after they entered. In the beginning, the chance of survival is perceived to be small, therefore firms invest little, thus making the low expectations of survival self-fulfilling.[24] Some firms, however, do achieve reasonable profits, so their increased chance of survival enables them to invest more. Other firms that cannot reach that profit level exit. Firms that survive the first periods continue to increase their investments until they reach sufficient profits to secure their position on the market. After that, they keep their investment steady or even decrease it for as long as their entry deterrence strategies remain effective. But then these non-increasing levels of investment reduce the chance of survival, all the more with entrants coming up with product innovations. As soon as the chance of survival drops below a certain level, the

[24] This probability of survival is described by a stochastic process.

incumbent facing the competition reconsiders its investment strategy and either increases investment again or leaves the market. Consistently with this description, simulations of the model show a high initial mortality rate (as a result of low investment initially), long lifespan for surviving firms (for as long as they can effectively deter entry), and 'coasting' states (slow decline) for very successful firms (those that decided to increase investment to try to remain competitive after they could no longer deter entry). In addition, the number of active firms is stable over time, which indicates that intensive exit accompanies intensive entry, resulting in a positive correlation between the two.

To summarise, active learning models allow us to predict the effects of entry costs on firm size, entry/exit and other variables through simulation, as shown in Table 4.4. A lower entry fee results in a less concentrated industry with a higher entry rate. Higher entry rate prompts incumbent firms to implement entry deterrence strategies, including innovation. Therefore, inasmuch as incumbent firms innovate in order to try and deter entry, lower entry barriers stimulate productivity growth through innovation. Furthermore, lower entry costs increase the correlation between entry and exit because the period during which unsuccessful entrants stay on the market becomes shorter.

Table 4.3 Active learning models: impact of changes in exogenous factors[a]

	Entry/exit rate	R&D	Firm productivity	Productivity of sector
Lower entry cost	+	+	+	+

[a] The comparative static analysis is based on simulations out of which the papers refer only to the effects of changes in entry costs. Therefore even though the other exogenous parameters play an important role in entry/exit, productivity and innovation, these relations are not explicitly revealed. These exogenous factors are fixed costs, efficacy of investment, market size, or demand.

Empirical verification

Active learning models are consistent with all the stylised facts that passive learning models purport to explain.[25] In addition, they explain how reallocation of resources among the incumbents may be the result of changes in the conditions regulating entry. Due to the nature of theoretical papers, these studies consider industry specific regulation. For instance, Olley and Pakes (1996) examine the US telecommunications equipment industry between 1974 and 1987 - the period coincided with the divestiture of AT&T and the beginning of market deregulation. They find increases in the rate of aggregate productivity growth after the market was deregulated. The reallocation of capital towards more productive establishments contributed primarily to these productivity increases. So, the two predictions stemming from the active learning models - higher productivity growth after reducing the costs of entry (through deregulation), and superior performance of the most productive firms - have been confirmed empirically. Furthermore, the positive correlation between entry and exit is also tested in

[25] For summary see Caves (1998) and Bartelsman and Doms (2000).

empirical studies. An example is Dunne et al. (1988) that analyses the US manufacturing industries over the period 1963-1982.

Conclusions

Active learning models aim at analysing firm-level and industry dynamics. These models are successful in explaining empirical findings, such as positive correlation between entry and exit, high mortality rate of entrants and the existence of a fraction of firms whose productivity remains high for a long time. The presented models consider changes of entry costs as an exogenous factor characterizing the industry. Under free entry active learning models find a high initial mortality rates and high entry rates, a reasonably long lifetime and an effective entry deterrence strategy for survivors, and eventually slow death for successful firms. When entry becomes less costly, competition intensifies, and even innovation increases a little. In the reviewed active learning models adjustment costs are not considered, but since these models provide a general framework to more specific analyses, the effects of adjustment costs could potentially be examined.

New developments in endogenous growth: entrepreneurship important to filter knowledge

Fitting in the endogenous growth theory, Audretsch et al. (2006) and Acs et al. (2004) developed a model that uses the concept of a filter between knowledge and economic relevant knowledge. The current endogenous growth theory does not completely clarify how or why spillovers actually arise. Therefore, their model identifies entrepreneurship as a mechanism that reduces this so-called knowledge filter. Stated otherwise, entrepreneurship is the transition mechanism converting knowledge into economically relevant knowledge.

In this perspective entrepreneurship is endogenous and depends on the entrepreneurial skills and investments in knowledge. The creation of a new firm is the endogenous reaction to investments in knowledge that incumbents have not completely appropriated. It is argued that additional entrepreneurial activity will increase as a rationale response to the creation of new knowledge. Hence, entrepreneurship is an important mechanism to facilitate the knowledge spillover, create new firms and make other ineffective firms obsolete, and eventually generate economic growth.

This new theory is still in its infancy and formalising the idea into both an input factor and a testable model is an area for further research. Besides measurement issues, such input factor is probably strongly related to human capital and hence difficult to disentangle in empirics. In addition, changes in ‘entrepreneurship’ are also the result of changes in entry and exit. As far as we are aware of, this new perspective on entrepreneurship does not look at the individual effects of entry and exit on the productivity performance.

4.4.3 Schumpeterian growth theory

In this section we consider two (interrelated) Schumpeterian growth models which focus on entry, exit, innovation and aggregate productivity in the economy.

Schumpeterian growth model A: sequence of ever more efficient temporary monopolists

The model[26]

Aghion and Howitt (1992 and 1997) present a model, which aims to provide an explanation for the stylised fact that firm entry and exit coincides with aggregate productivity growth. Stated otherwise, they introduce Schumpeter's idea of creative destruction combined with innovations which improve the quality of products (i.e. the so-called quality ladder or vertical innovations) in their model. According to their model, entrants bring innovations to the market and replace the incumbents, as a result of their more efficient production technology. Better products replace obsolete previous ones. This replacement leads to productivity growth at the aggregate level.

The model describes an economy consisting of two sectors, one producing innovations and the other final consumption goods, using innovations as one of the inputs. The innovation industry uses (skilled and specialised) labour as input. Moreover, the number of innovations is proportional to the accumulated knowledge, which implies increasing returns to scale due to exploitation of knowledge spillovers. Also, there are two exogenous parameters: one is the probability of a successful innovation and the other is the 'productivity step'. The latter is the difference in productivity between the old and the new technologies applied in the consumption good industry. There is no entry and exit in the innovation industry.

All consumption goods are homogenous and produced by a monopolist firm. The consumption good industry also uses (skilled) labour as an input, but, unlike in the innovation industry, returns to scale here are decreasing. Decreasing returns to scale (or increasing marginal costs) are needed to balance the marginal labour productivity levels between the innovation and the consumption goods, thus ruling out a hypothetical situation when the whole labour force is committed to one sector or the other. In fact, the extent of decreasing returns to scale is another exogenous parameter in the model. The labour force is fixed, and employees in the innovation industry cannot make consumption goods. Therefore, there is a trade-off between higher productivity growth tomorrow and more consumption today.

The productivity of the consumption good industry increases with firm entry. Entrants buy innovations from the innovation industry and apply them in the consumption good industry as process innovations. The price an entrant pays for the innovation is its investment that it compares to the discounted profits over a lifetime. The discount factor is time preference. So, the lower the time preference, the more the entrant is willing to pay for an innovation, and the

[26] Section 4.4.3 is based on memos that are available on request. In contrast to the original paper of Aghion and Howitt (1992), section 4.4.3.1 distinguishes only the innovation industry and the consumption good industry, while it ignores the intermediate sector. We have done this to simplify the explanation while it does not harm the arguments. The reason is that a skilled employee in the intermediate sector produces one unit of intermediates (AH, 1992, equation 2.2). Moreover, the number of intermediates is the input in the production function of consumption goods. Hence this equals employment in the intermediate good sector. Therefore, the intermediate good sector does not need to be mentioned for the explanation.

more innovations are produced. The monopolist in the consumption goods industry does not innovate because innovation will reduce its rents. The entrant, on the other hand, has a strong incentive to innovate in order to replace the less efficient incumbent enjoying a monopolist's rent. Each time a new firm enters the consumption good industry, it forces the actual incumbent monopolist to exit. In the course of time the consumption good industry is populated by a sequence of ever more productive temporary monopolists. The speed of this replacement determines productivity growth. Entry in the consumption good industry is endogenous because it depends on the price of innovation 'goods'. This price depends on a number of exogenous factors that may be affected through policy interventions, as we discuss in the next chapter.

The fundamental determinant of aggregate productivity growth is the division of (skilled) labour force between the two sectors. In the equilibrium, this division is determined by equal marginal labour productivity in both industries. If the economy adjusts from one equilibrium to another after a shock, labour moves from one industry to the other until the marginal product of labour in both industries are equal again. Decreasing returns to scale in the consumption good industry is the 'bottleneck' required to ensure the equilibrium and to prevent the explosive growth on the back of increasing returns to scale in the innovation industry. The allocation of the fixed labour force over the industries, and hence productivity and entry/exit, is determined by the extent of decreasing returns to scale in the consumption good industry, the probability of a successful innovation, and the productivity step. The workings of these exogenous factors are discussed in more detail in the following section. The costs of adjustment are ignored.

Results and exogenous factors

Table 4.5 indicates the signs of the impacts of a change in the exogenous factors on the key variables in the Aghion and Howitt model.

Table 4.4 Schumpeterian growth model A: impacts of changes in exogenous factors

	Entry rate = exit rate	Productivity growth economy
Larger size economy	+	+
Probability parameter of innovation success	++	++
Larger productivity step	+	++
Less decreasing returns to scale in consumption goods sector	–	–
Higher time preference	–	–

If the size of the economy (i.e. population) increases, the additional employment is equally distributed among the industries. With more people working in the innovations sector, and more innovations produced, productivity growth and consumption increase at same rate.

If the probability of innovation success rises there are two positive effects on having a successful innovation, thus higher productivity. First, employment in the innovation industry rises and employment in the consumption good industry declines with the same amount. The reason is that the innovation industry has become relatively more productive. Second, each employee in the innovation industry has become more productive. The interaction of more employment in the innovation industry and the higher productivity in that industry produces the higher number of innovations in a year, and consequently the number of entrants and exits in the consumption good industry. The productivity in the latter industry increases with the number of entrants multiplied by the size of the exogenous productivity step. There is a negative effect as well as a an increase in the probability of innovation success will also increase the rate of destruction. However, the positive effects dominate.

If the productivity step increases, the value of innovations increases. Therefore, the wages in the innovation industry increase and workers move from the consumption good industry to the innovation industry. Employment in the innovation industry increases with the same amount of the decline in the consumption goods industry. The number of entries and exits in the consumption good industry in a period rise proportionally. Productivity growth accelerates more as each innovation causes a productivity increase with an innovation step.

Less decreasing returns to scale in the consumption good industry makes employees in this industry more productive. Therefore, employees move from the innovation industry to the consumption goods industry. Consequently, consumption increases, but the entry and exit rates and the productivity growth fall because fewer innovations are now produced.

A higher time preference (a shorter time horizon) reduces demand for innovations from the potential entrants because the expected profits from innovation are discounted against a higher rate. Lower demand for innovation leads to less employment in the innovation industry, fewer innovations produced and lower productivity growth.

Empirical verification

The model has, to our knowledge, not been empirically verified yet. Still, it supports some stylised facts. First, the model defines the stylised fact that learning from other people increases your own productivity in the production function of the innovation industry. There the exploitation of the accumulated knowledge determines the labour productivity of the employees in that industry, which is another way of saying that these employees learn. In its turn, the accumulated knowledge is another way of saying what other people know. The model relates to the empiricism in chapter 3 in the following way. In the end, there is 100% turnover in the model because each time a firm enters the consumption good industry it forces the incumbent monopolist to exit. In contrast, the incumbent does not innovate and hence does not contribute to aggregate productivity growth. This implication contradicts with the evidence that part of the observed aggregate productivity growth is attributable to incumbent firms.

Conclusion

The main exogenous factors are: size of the economy (total population), probability of a successful innovation, extent of decreasing returns to scale in the consumption good industry, productivity step between two consecutive innovations, and time preference. The endogenous

variables are: spending on innovation, entry (=exit), and aggregate productivity growth. The model derives results for the entire economy. The adjustment costs are ignored.
As the model distinguishes appropriability, intertemporal spillover and business stealing effects, a welfare analysis is possible comparing the economy's growth rate under the laissez-faire situation with the optimal case. The business stealing effect provides too little research under a social planner, whereas the two other effects provide too little research under laissez-faire.

Schumpeterian growth model B: high-productivity entrants motivate incumbents to innovate

The model
Following up on the fundamentals of earlier work of Aghion and Howitt, the idea has been postulated that the distance to the technological frontier matters for countries or industries. Aghion et al. (2004, 2006) developed a model in which entry threat has an impact on incumbents' innovation expenditures and, consequently, productivity growth. In the model, entry threat is an exogenous parameter which measures the probability that a (foreign) firm enters the (home) market. This impact on innovation is non-uniform across firms and industries. Higher threat of entry leads to higher innovation expenditures and higher productivity growth of incumbents, which are already highly efficient. It discourages less efficient incumbents to spend on innovation. A similar mechanism works at the industry level: world leaders have a stronger incentive to innovate than world laggards. The reason for this heterogeneity in the incumbents' response to entry threat is simple: while the costs of innovation are the same for all firms, the market leaders have a higher chance of retaining their leadership in the face of entry than the laggards have of gaining it.

The model describes an economy with different industries each producing a distinct good. If all firms in a given industry have access to the same technology, competition drives their profits down to zero. On the other hand, if production technologies differ, only the firm with the most efficient technology (leader) ends up producing, and receives a non-zero profit. The possibility of receiving a profit through better technology gives firms an incentive to invest in innovation. However, expected profits to firms from a successful innovation differ depending on the distance of the industry in question to the *technological frontier* (i.e., the technology giving the highest possible level of output given the inputs), and the threat of entry, measured as the probability of a new firm entering that industry. The authors consider three types of firms: type 1: industry leaders with the level of technology on the frontier; type 2: industry leaders with the level of technology behind the frontier; and type 3: firms in an industry without a leader. The entrants must always have the frontier technology, otherwise they will not be able to recoup their entry costs as profits are non-zero only for the leader.

Clearly, as the threat of entry increases, the incentive to innovate is higher for type 1 firms than for type 2, because type 2 firms are at a disadvantage of having an inferior technology compared to the prospective entrant. Some additional assumptions of the model make this conclusion even stronger: in fact, the threat of entry increases innovation expenditures for type 1 firms, and reduces for type 2 firms. Type 3 firms never innovate because they are too far from the frontier to earn profits from innovation. Thus the model predicts different responses to the threat of entry of incumbent firms and hence different consequences for their contribution to aggregate productivity, depending on how technologically advanced they are.

Results and exogenous factors

The main exogenous factors are the threat of entry, the size of a productivity step (the difference in productivity between type 1 and type 2, and type 2 and type 3 firms) and the substitution elasticity between the product varieties which are supplied by the industry. This elasticity is the only determinant of the profit rate, which therefore is indirectly exogenous. The profit rate can also be described as a fixed share of the productivity level of each firm. Hence low productive firms earn less then high productive firms.

Table 4.6 summarises the main implications of the model for the two variables of our interest - innovation expenditures and aggregate productivity. Despite different effects on innovation spending depending on the type of the firm (see above), higher threat of entry always leads to aggregate productivity growth. It is likely to be endogenous for the industries with type 1 firms, where the growth comes from internal efforts to improve technology, and exogenous for the industries with type 2 and 3 firms, where productivity will grow due to entrants bringing more productive technologies from outside.

Table 4.5 Schumpeterian growth model B: impacts changes in exogenous factors

	Innovation expenditures		Productivity industry
	Type 1 (high productive) incumbents	Type 2 (medium productive) incumbents	
Entry threat increases	+	−	+
Productivity step increases	++	+	+
Leader’s profit	+	+	+

A larger productivity step means more radical and profitable innovations. Therefore, holding the threat of entry fixed, this factor has a positive effect on innovation spending and aggregate productivity growth in all cases. It is stronger for more technologically advanced firms (type 1) because there is a larger chance for them to reap the benefits from innovation, whether the probability of entry is high or low, whereas type 2 firms will benefit if and only if there is no entry. Similarly, higher profits from being an industry leader make leadership from innovation more attractive, as compared to the zero profit, no innovation situation of type 3 firms. So, higher profits increase innovation spending of every type of leader, and because more innovation leads to better technology, aggregate productivity will grow as well.

Empirical verification

Aghion et al. (2004 and 2006) support their theoretical findings with firm-level data from the UK, using foreign direct investment as the measure of entry threat. The model relates to the empiricism in chapter 3 in two respects. It uses the stylised fact that entry and exit go hand in hand, and it explains why aggregate productivity growth is partly due to changes in market shares of incumbents each operating at different levels of productivity. An explicit link to firm heterogeneity, reflecting a salient empirical fact, is another enviable feature of this model.

Conclusions

The model analyses an economy consisting of industry sectors populated by firms of three types: type 1, an industry monopolist with a frontier technology; type 2, an industry monopolist with a sub-frontier technology; and type 3, a competitive firm with a still lower level of technology accessible to all in a given industry. All industries experience a threat of entry of a new firm with frontier technology that would replace the less productive incumbent(s), but only type 1 and 2 firms innovate in an attempt to improve their technologies. As the threat of entry increases, type 1 firms innovate more, and type 2 firms innovate less. Aggregate productivity increases with entry for every type of firm, but actual entry is not required to encourage type 1 firms to invest more in innovation and thus increase aggregate productivity.

Investment strategies of world leaders and laggards

Discrepancies in efficiency levels are also observed across countries and industries. In addition, the same economy can be the world leader in one industry, while be a laggard in another industry. Therefore, the position of the economy or an industry vis-à-vis the world's technological frontier may be also an important factor for policy.

According to Acemoglu et al. (2002), high-skill managers are more important for innovation than for adoption. Therefore, selection of the managers plays little role in backward economies, while the role of manager selection increases more and more when the economy approaches the frontier. Based on this, relatively backward economies are characterised by the so called 'investment-based strategies', with long-term relationships between firms and managers, high average size and age of firms, large average investments, but little selection of managers. But as the economy approaches the frontier, there is a switch to 'innovation-based strategies', characterised by short-term relationships, younger firms, less investment and better selection of managers. As a result, the rates of entry and exit increase when the economy approaches the technological frontier and switches from adoption to innovation.

With respect to exit and entry, it means that relatively backward countries should pursue investment strategies (i.e. provide subsidies and limit competition, which will also slow down entry/exit),[27] and countries that are close to the frontier should generally encourage innovation by competition.

---

[27] In particular, Acemoglu et al. (2002) show that if the switch from adaptation to innovation occurs too soon, policies encouraging the investment-based strategy are indeed beneficial. However, they also stress that if switch is desirable, but the society cannot realise the switch, it may fall into a 'non-convergence trap'.

4.5 Summary

In this section, we draw conclusions from the literature concerning the interaction between firm entry and exit and productivity. Table 4.7 summarises the important facts about the three classes of theoretical models delineated in figure 4.1 and it elaborates on the four questions raised in the introduction to this chapter.

Exogenous factors affecting entry/exit and productivity

The key differences among the three groups of models that determine the scope of entry/exit for productivity growth are:

- The nature of productivity changes;
- The type of the market environment that firms operate in.

Table 4.6 Summary of theoretical assumptions and predictions of reviewed theories

	Static efficiency	Exogenous growth	Endogenous growth
Exogenous factors affecting entry/exit and productivity			
Nature of productivity changes	Better allocation and utilisation of the existing technology and resources	Exogenous technological change (e.g. from abroad, or from another industry)	Investment in innovations (product or process)
Exogenous factors[a]	Technology	Parameters of cost function (inc. entry cost) Parameters of demand, size of productivity change	Parameters of cost function (inc. entry cost) Parameters of demand Ability, probability, type and step of innovation, gains of innovation
Role of entry/exit in determining productivity			
Relation between P and EX	EX⇒P	P⇒EX	EX⇔P
EX	Exogenous	Endogenous	Endogenous
P	Endogenous	Exogenous	Endogenous
Effect of entry threat on incumbents	It stimulates efficiency of incumbents	No relation	Positive effect on innovations in advanced sectors; but negative in other sectors
	Firm/industry	Firm/industry	Firm/industry and aggregate[b]
Costs corresponding to firm entry/exit			
Entry cost	Implicitly present	Present	Present
Exit cost	Ignored	Ignored	Ignored
Adjustment cost (reallocation effects)	Ignored	Present in passive learning models	Typically ignored[c]

[a] Differs per model.
[b] Only in some models.
[c] In some models adjustment costs are included as cost of increasing volume.

Theories agree on the effect of some factors (at least to the extent to which they incorporate these factors and draw conclusions with respect to their effects), for example entry costs, on entry and exit and productivity: lower entry costs lead to more entry and higher productivity.[28] However, the effects of changes in other factors, such as fixed costs, some demand parameters, or the ability to develop an effective innovation, may vary.

The role of entry and exit in determining productivity

The role of entry and exit in determining productivity depends on the model. As long as we are concerned only with static efficiency, technology does not change. Efficient entry increases aggregate productivity directly, since these firms have higher than average productivity and it

[28] As we explained in chapter 2, higher productivity does not always mean higher welfare. Note that only the Schumpeter growth model A offers the opportunity for a welfare analysis.

shifts market shares from inefficient towards efficient firms that are more productive. In addition to entry, the *threat of entry* may also play a role. In static models it increases incentives for incumbents to eliminate their X-inefficiencies, thus increasing their individual productivity, and therefore the industry's aggregate productivity. Therefore, in static models, entry/exit of firms is a mechanism through which competition improves productivity by reallocating production towards more efficient firms and creating competitive pressure on inefficient incumbents.

Dynamic efficiency changes arise if the level of technology changes. In this case, we also need to take into account the nature of productivity change. When exogenous productivity developments influence an industry, innovations fall from the sky as manna and affect productivity, which in turn affects entry and exit decisions. Since productivity change is exogenous, it is not affected by entry threat.

In contrast, in the models that endogenise productivity change through investments made by individual firms, innovations affect both entry/exit decisions and productivity. As we learn from the Schumpeterian growth model B, the threat of entry may also play an important role in determining innovation expenditures and thus productivity. It may increase the incumbent's incentives to innovate for as long as their outcome keeps entrants out. However, if the chance to outperform a potentially more productive entrant is small, the threat of entry suppresses the innovation activity of the incumbent. The overall effect on innovation spending by incumbents can be positive or negative, depending on country/industry, but the effect of entry threat on aggregate productivity is always positive.

Aggregation level

In order to consider policy options, it is also important to understand at which aggregation level the theory addresses productivity: firm, industry or macro-economic level. Most theories focus on the industry level, and do not take into account the effects of reallocation in the economy as a whole.

Costs corresponding to firm entry and exit

Theoretical models generally include costs of entry, while the costs of adjustment or exit are typically ignored.[29] These costs may be considerable in practice, especially in industries where labour input is a large part of total costs or where assets are specific and cannot easily be re-employed in other production processes. Adjustment costs are likely to be an important issue in labour-intensive industries because labour is often protected and is therefore expensive to shed. Potential entrants that want to enter the market do take adjustment costs (discounted by the probability of having to exit the market) and exit costs into account, and therefore the direct costs of entry are only part of the total costs that determine whether or not a firm enter the

[29] Here we mean specifically the costs of downward adjustment as opposed to the costs of capacity expansion which are sometimes included, e.g. in the life-cycle models.

market. Because adjustment costs enter separately, their omission in theories is often a welcome simplification. It does not affect the relationships between the costs of entry, productivity and other variables. However, policy makers should take these costs into account, as we discuss in the next chapter.

5 Entry related policies promoting productivity

5.1 Introduction

This chapter derives implications for policies from the theories and empirics reviewed in chapter 4 and 3 respectively. Both preceding chapters bear one main message - that free entry appears to be conducive to productivity growth under a wide range of circumstances. This suggests a high potential for policy interventions in the areas affecting industry dynamics, such as reducing barriers to entry and streamlining exit procedures. Both chapters also suggests the importance of measures addressing both entry and exit being parts of the same policy package, as potential entrants factor the difficulty of exiting the market in their decisions to enter. Hence, at the outset, there seems to be a case for a universal policy to stimulate productivity growth through more liberal entry. In addition, differences in productivity growth contributions of firm entry and exit between industries reveal a pattern consistent with the idea of firm entry facilitating technological innovation. Consequently, it is also useful to discuss the scope for industry-specific policy as opposed to general policies.

To think about the role of policy, we apply a welfare economic approach. This approach assumes that the results of markets are an efficient outcome from a welfare perspective as long as no market and government failures exist. In that respect, we underline two elements before embarking into details.

First, in general, increasing productivity as indicator for welfare should be the main aim of policy and not the number of entrants or firms, the intensity of competition or the extent of innovations. The latter are only means or determinants to improve productivity. For instance, the models in chapter 4 show that entry and innovation themselves can be endogenous.

Second, welfare is not optimal if market failures or government failure are present. If market failures (i.e. situations where private actions do not lead to maximum total welfare) exist, then government may consider intervening if and only if government failures are smaller than those market failures; i.e. if the social benefits of intervention exceed the costs. Policies that trigger those above mentioned determinants of productivity are only effective if those determinants are exogenous for the actors in the market, because then they tackle the core problem of market failure at its roots.

Hence, we begin this chapter by outlining the market failures concerning (the costs of firm's) entry, and then proceed to discussing policy options to repair those failures. As discussed, the existence of a potential market failure does not a priori imply that government intervention is needed. Whether such policy measure actually improves social welfare depends also on the extent of government failures. Therefore, an ex ante cost-benefit analysis needs to precede intervention. That kind of analysis is beyond the scope of the document.

Henceforth, we entirely focus on policy options to promote (productivity) growth through firm entry. We do not discuss the policy options related to other exogenous factors that determine productivity such as the abilities of workers (and other human capital issues), non-domestic spillover effects of accumulated knowledge and foreign technological progress after a successful innovation.[30] Moreover, the reader should keep in mind that for firms operating in the Netherlands being an open but small economy, most of the new ideas or, stated otherwise, productivity growth come from abroad. Discussing those other factors is also out of the scope of this document and already extensively considered in earlier CPB-studies (e.g. Canton et al., 2005 and Cornet et al., 2006).

5.2 Market failures concerning entry and policy options

As discussed in chapter 2, entrants bear costs before entry. For example, they need to invest in plants or equipment, buy licences and permits, and spend on finding workers and advertising. Applying the welfare approach, one could argue that there are no market failures as long as entrants pay market prices for these subjects. However, those costs as such may already lead to incomplete markets implying that market do not deliver the efficient allocation mechanism. Chapter 2 distinguishes three general types of entry barriers, i.e. technological barriers, strategic barriers and institutional barriers. Particularly, strategic barriers and institutional barriers affect the mechanisms of a market. The main question to be addressed is whether market failures related to those barriers prevent firms from entering the market or being a threat for incumbents hampering their incentives to innovate and generate more welfare. In this respect, we outline three market failures and pose policy options:[31]

- Market power of incumbents including relevance of institutional barriers;
- Capital market failures inducing capital rationing for new businesses;
- Production externalities (see 5.3).

5.2.1 Market power and capital market failures

Market power: expected entry costs higher due to abuse of market power of incumbents
Strategic barriers may provide incumbents market power. The costs that new firms bear due to deliberate incumbents' actions to prevent the entry of new (and potentially more efficient) firms imply a market failure because of imperfect competition. Such entry barriers limit the threat of

[30] Note that abilities and other human capital issues are also related to the quality of new entrepreneurs. Improvements in the capabilities of entrepreneurs will positively affect the potentials for productivity developments.
[31] Market failures related to network externalities are discussed in section 5.5 as those market failures are closely linked to particular industries. Market failures related to asymmetric information are mainly discussed as part of the capital market failure. We ignore the problem of asymmetric information in the case of lack of social capital (i.e. network contacts, social relationships) of entrants.

entry for incumbents. Incumbents then can charge relatively higher prices above marginal costs. This generates static inefficiencies. Moreover, an incumbent, particularly a monopolist, has no incentive to innovate if there is no entry threat, since by innovating she replaces her own rents (replacement effect, Arrow 1962).

Theory mentions deliberate creation of overcapacity or limit pricing (price setting just below the marginal costs of potential entrants), cartels or tying and bundling of products (goods are only valuable to consumers when used together) as examples of entry deterrence. In practice, Bunch and Smiley (1992) find that these instruments of entry deterrence are hardly used. Instead, more frequently used are advertising, filling market niches to reduce unmet demand, and suppressing information on profitability to reduce the attractiveness of business to potential entrants (see also Chang and Tang, 2001).

Adequate competition legislation should repair this market failure if market power of (dominant) incumbents gives reason to suspect the abuse of power. In the Netherlands, competition authorities such as the NMa or the Directorate General for Competition of the European Commission carry out competition policy. For instance, the NMa enforces fair competition in all sectors of the Dutch economy and it takes action against parties who participate in a cartel, for instance by fixing prices. The NMa also acts against parties who abuse a dominant position and it assesses mergers and acquisitions.[32]

Market power: entry costs higher due to institutional barriers

Similarly, institutional barriers may provide incumbents with market power. Potential entrants face administrative costs, such as acquiring permits, as a condition to start operations. Examples include building permits, environmental policy or food safety approvals. Because entrants are relatively small, the burden of compulsory administrative costs is heavier for them than for incumbents, as these costs are not related to production volume. As any barriers to entry including those for foreign firms, administrative ‘red tape’ may reduce productivity by limiting competition.

In contrast to strategic barriers, the case for reducing institutional barriers to entry is not at all universal. Costs related to institutional barriers per se do not necessarily imply a constraint in minimising the sum of market and government failures. Instead, the controls financed by these costs are meant to prevent market failures due to information asymmetry or missing of market mechanisms. Similarly, they exercise controls when it comes to health and safety, quality control, environmental protection etc. The latter is related to the broader concept of welfare, being not only productivity but also health and environmental issues, and even equity issues (i.e. redistribution). As we argued in chapter 2, maximum productivity does not correspond to maximum welfare. For example, the gains from choosing not to implement the prescribed

[32] In practice, with respect to deter entry, this is not an easy task as for potential entrants it is hard to complain when they have not started at all.

health and safety measures at work accrue to the firm alone, but the costs are spread among the many taxpayers (negative externalities). So, while it is optimal for the firm not to implement health and safety issues, this is clearly not welfare-maximising for the society at large.

Also against the unconditional lowering of institutional entry barriers is the point made in Bunch and Smiley (1992) that such policies may inadvertently raise another market failure, as the probability of anticompetitive conduct with entry deterrence of incumbents may increase. If, however, this should be the case, there might be a case for competition authorities if incumbents abuse their market power.

Empirical evidence, as Nicoletti and Scarpetta (2003) showed for the OECD countries, suggests that a reduction in the institutional entry barriers leads to more entry and higher aggregate productivity. In the Netherlands a number of institutional rigidities were relaxed during 1994-2004 as they were regarded redundant in the 'Marktwerking, Deregulering en Wetgevingskwaliteit (MDW)' directive.[33] In that respect, it is interesting to note that as chapter 3 showed, PMR is still relatively high in the Netherlands compared with other OECD-countries.

All in all, it is safe to conclude that, although reducing institutional entry barriers seems a promising policy direction, the benefits it brings to productivity should be assessed against potential welfare losses it may incur.

Capital market failure due to capital constraints

Financial markets are an important selection mechanism for take-off of entrants on product markets. Having often too less money to enter the market, potential (small) entrants need funds (i.e. bank lending, venture capital, stock markets, or private equity) to pay the entry costs before they can recoup them later. Frequently they have difficulties getting those funds.

Capital constraints might limit entry because of two main reasons. One is higher risk inherent in financing a new business than that of an incumbent. This is, however, not a market failure. The other reason is a capital market failure due to information asymmetry between the supplier of funds (e.g. banks) and potential entrants. The lack of collateral reputation and the absence of track records exacerbate the information asymmetry. The result of this capital market failure is that it can generate less entry than optimal, and lowers aggregate productivity growth.

The potential relevance of pre-market selection mechanism of financial markets can be empirically illustrated as follows. Chapter 3 shows significantly larger variations in productivity levels of new firms and higher shares of employment of entrants in total employment in the US than in Europe. A possible reason for this difference is that in the US there is less pre-market selection and more selection by actual customers than in Europe. Differences in providing

[33] See the final progress report Ministry of Economic Affairs (2003).

capital between the US and Europe are an interesting avenue for further research, particularly given the current word crises as result of the functioning of the financial markets.

Empirical literature confirms that sufficient start-up capital is a critical constraint for a large majority of entrants, especially in the technologically advanced industries. Lerner (1996) observes that the presence of venture capital financing stimulates growth in high-tech firms, and Hellmann and Puri (2000) give evidence that it can decrease the period within which new products are introduced. Hall (2005) finds that market imperfections limit growth of 'early-stage' firms. Boot and Schmeits (2004) evaluate a number of capital market policies in the Netherlands towards innovative enterprises and conclude that there are a couple of bottlenecks. They derive the following criteria of an effective policy to address capital market failures. First, both investors and borrowers should share risks. A way of achieving this is equity financing, where the borrower and the investor share both the upside benefits and the downside risks of the investment. Second, the investors should carry out the regulation. Third, the amount of the loan must be between 1 and 4 millions of euros.

Hence, policy should address capital market failures if they are present, but this policy needs substantial information to be conducted effectively. In particular, it is important to understand to what extent asymmetric information (which is a market failure) and a higher risk of investing in new ventures (which is not) cause credit constraints including external equity.

Is subsidising entry welfare enhancing?[34]

A subsidy given to each entrant seems a simple policy to repair whatever legitimate market failure concerning entry. For instance, one could claim that entrants embody new technologies or produce new goods and services all contributing to higher welfare. Indeed, entry may generate a change in dynamic efficiency. One could therefore argue that potential entrants should be stimulated to enter the market. This policy of entry subsidies, however, is neither legitimised nor efficient because the market failure is not evident. There are three reasons. First, indiscriminate subsidising of entrants will stimulate the emergence of 'revolving door firms', which deliberately enter in order to get the subsidy and exit thereafter. Or it supports an inefficient firm to survive.[35] Hence, such a policy distorts both market selection and the learning process of 'bad' entrants of their inefficiency. Consequently, it delays the exit of those firms. Second, this policy is not efficient because efficient firms that would have entered anyway end up receiving subsidies. Finally, it is also not efficient because subsidies, paid with tax money,

[34] Another policy option could be that policy aims to improve the pre-market selection of entrants through interviews and checking business plans. This option requires a lot of detailed information and knowledge of government. In fact, government failures can result from information asymmetry or the principal-agent problem. For instance, in industries where technology is moving fast, governments face large difficulties permanently obtaining all the information needed for taking appropriate decisions.

[35] See eg. Santarelli and Vivarelli, 2006, pp. 24-25.

withdraw scarce resources from other productive activities, which in its turn will have a negative impact on productivity of the country. In other words, it creates a deadweight loss.

Taxes and social security expenditures

The main text primarily focuses on policies related to influence the entry and exit costs, but other policies may affect the decision to enter the market as well (see also section 5.4). As already referred to in chapter 2, taxes on profits affect the opportunities for entry as high taxes lower the expected value of profits of entrants. Hence, the tax system may distort the incentives for entry including the threat of entry.

Similarly, there can be a link with the social security system via the development of employment. Lowering the discrepancy between gross and net wages, and increasing the replacement rate (the net unemployment compensation as a percentage of the net wage earnings) have a negative effect on the (natural) rate of unemployment increasing the employment. All else equal, this will positively influence the number of entrants.

5.2.2 Entry, market power and trade-off between static and dynamic efficiency

According to theory, more entry stimulates competition between firms and improves static efficiency through reducing the deadweight loss. Entry can also generate higher dynamic efficiency as entrants can embody new products. Theory is, however, less clear-cut about the relationship between competition and innovation: it can either be positive or negative. The implications for policy get even more complicated if the relationship between competition and innovation turns out to be an inverted U-curve (see Aghion et al., 2005a).[36] Then policies that encourage competition via entry may be detrimental for innovation beyond some level of competition.

What does it mean for policy if an inverted-U relationship actually exists? Is there some kind of optimal threshold of entry barriers stimulating competition, taking account of issues related to, amongst others, health and safety? Again, we need to be certain of the type of market failure here. We state that the market failure is not evident a priori. Let us dig deeper.

First and utmost we should distinguish between ex post and ex ante competition. Ex post, innovating firms need some market power to limit free access to their innovation, whereas this market power is not essential in case of ex ante competition. New firms will only enter if they can recoup their spending on entry including innovation expenditures. Competitors should therefore not have free access to their innovation and imitate their idea costless after entering the market. So, those entrants need market power via an intellectual property right (IPR) or a mark-up to generate positive profits after entering the market. Hence, ex post competition in industries where innovative firms operate has to be imperfect (see also models of Aghion and Howitt in chapter 4). In contrast, competition can ex ante be perfect with each firm pricing at marginal costs and still provide enough incentives to innovate (via escape competition). Hence,

[36] More intense competition increases the incentive for firms to innovate to escape competition but it reduces the incentive for lagging firms to catch up. But if competition becomes tougher, the reduction in the catch up effect gradually starts to dominate the escape effect.

in principle, policy can lower institutional entry barriers if the IPR is properly organised including taking account of trade-offs (see box IPR).

Second, innovations in heterogeneous products (or product diversification) may lead to new niches in the market reducing measured competition.[37] As long as there is no abuse of market power or institutional barriers in those niches, and this product diversification reflects consumer preferences, there is no market failure.

Finally, note again the importance for policy to focus on productivity and welfare instead of the intensity of competition or the amount of innovation expenditures. According to the theory of Aghion et al. (2006) discussed in section 4.4.3, a decline in innovation expenditures (of incumbents) due to more intense competition in an industry can go hand in hand with higher productivity of that particular industry. The reason is that the entry of a foreign leader replaces a less productive domestic firm. The latter has no incentive to imitate or innovate after more intense competition due to the large productivity gap.

Intellectual Property Rights (IPR)

IPR might work as a temporary entry barrier applied selectively to imitators that enable innovating firms to earn a mark-up on the new product. It gives the owner the right to exclude others from the innovation and the right to charge others a price if the innovation is licensed. IPR are less related to firm entry itself. In practice, firms possess many patents and also within firms there is creation and destruction of products.

The IPRs are not meant to reduce entry as such. Because the royalties are applied selectively, to discourage firms from imitation and encourage innovation, a policy to protect intellectual property is consistent with the generic policy to reduce barriers to entry discussed earlier.

Relevant to policy is the debate on the scale of optimal protection. The IPR creates ex post market power for the inventor. A higher value of IPR stimulates innovation ex ante but generates a higher deadweight loss as well ex post. Hence, a trade off occurs between stimulating innovation and reducing the deadweight loss in the design of IPR. Then, the main issue is where to stop with protection or the duration of protection. The issue is also related to the social welfare dilemma between knowledge creation and knowledge diffusion (see CPB, 2002).

Looking at the empirical evidence, competition seems at least to encourage static efficiency. Often, in practice, increased competition lead to increased investment in R&D and innovation and that lead to faster productivity growth (see CPB 2008 for an overview). Evidence for an inverted U-curve is scarce. Aghion et al. (2005a) find support for such relationship for the British manufacturing industries. In contrast, Creusen et al. (2006) come up with a positive relationship between competition and innovation for the Dutch retail trade. Moreover, van der Wiel et al. (2008) show that the effect of increasing competition is positive for firms close to the technological frontier as well as for firms further away from the frontier. Currently, the empirical evidence seems to be in favour for more intense competition (e.g. through entry) leading to higher static and dynamic efficiency.

[37] Note that this action can produce the inverted U-curve in practice, despite that competition may always stimulate innovation.

5.3 Production externalities related to entry

Another market failure to be discussed are production externalities. Production externalities with respect to entry arise if activities of entrants generate benefits or costs that the entrant cannot (or do not) take into account.

Positive production externalities

An example of positive externalities is knowledge spillovers. R&D or innovation activities generate knowledge spillover effects creating welfare benefits that cannot be fully appropriated by the inventor. For instance, discussed in the models by Aghion and Howitt (1992) in chapter 4, the innovation industry produces innovations that other industries apply then, perpetuating productivity growth in the economy.[38]

Empirical evidence is generally supportive of innovation or knowledge spillovers. It appears, for instance, that spending on innovation abroad contributes positively to productivity in the home country.[39] In addition, there is empirical research that suggests that international exchange of students leads to higher exploitation of the positive external effects of accumulated knowledge.[40]

Innovation spillovers lead to a market failure of too little innovation as the social rate of return is larger than the private rate of return. Because of positive innovation spillovers, a higher mark-up after the innovation may be socially optimal (than the market outcome).

The implications for policy are, however, not automatically that entry should be encouraged less. Rather, subsidies on research and development should accompany entry-stimulating policies to achieve a fuller effect on productivity. Those subsidies tackle the core problem of the market failure here: the appropriability effect and the intertemporal spillover effect.
The R&D tax credit program WBSO is an example of such subsidy in the Netherlands. This subsidy was implemented to stimulate private R&D-expenditures in 1994.[41] Since then, it has been changed frequently. For instance, there were indications for market failures for innovative starters. In 2001, an additional tax credit was offered to new and innovating firms: the starters program. It turned out that this starters program yields extra R&D-labour activity generating knowledge (see Cornet and Vroomen, 2005). The current cabinet aims to intensify the WBSO and the innovation credits to further stimulate innovation.

[38] See also Grossman and Helpman (1991).

[39] See e.g. Coe and Helpman (1995), Coe et al. (1997) and Jacobs et al. (2002).

[40] See Park (2004).

[41] Recently, commissioned by the Netherlands Ministry of Economic Affairs, EIM (2007) concluded that the WBSO is a properly functioning regulation that encourages private R&D expenditure. Moreover, the impact is greatest for small firms and the additional investment (i.e. additional funds from firms' own resources) becomes lower as size increases.

Negative production externalities: business stealing

Most theoretical models in chapter 4 assume that entrants replace exiting incumbents. The potential market failure in this case might be the occurrence of business stealing by entrants.

Mankiw and Whinston (1986) show that in an imperfectly competitive market of homogenous goods where identical firms with fixed costs compete imperfectly, business stealing may lead to too much entry (see also Aghion and Howitt (1992) in chapter 4). Innovative entrants base their entry decision on their own expected profits after entry and do not take into account the losses to the incumbent firms due to their technology becoming obsolete. In other words, ceteris paribus the effect of innovation, the private marginal costs and benefits of entry are not equal to the social costs and benefits, which gives rise to welfare implications of business stealing. In that case, market entry restrictions can be socially required.

Policy implications

Either depending on whether the positive or negative effect of production externalities dominates, more entry can increase or decrease (dynamic) welfare. Too few firms enter the market if they are not able to appropriate the rise in consumer surplus due to their entry. On the other hand, there can be too many firms entering the market.

The empirical evidence suggests that the positive effect of knowledge spillovers is larger than the negative effect of business stealing. In fact, Mankiw and Whinston (1986) show that increasing the number of firms until the profits are zero is welfare maximising if firms act as price-takers. Hence, the general lesson for policy is, again, that entry should be encouraged to bring actual entry closer to the socially optimal. In the Netherlands, policies such as IPR and WBSO already consider production externalities.

5.4 Related policies to adjustment and exit costs

Chapters 3 and 4 already put forward the importance of the costs of adjustment and exit costs in the decision of a firm to enter the market. Industry's productivity is being improved if efficient firms gain market shares relative to inefficient firms, particularly if the latter are forced to leave the market. The emerging firms and industries need workers who must be withdrawn from the less efficient firms and industries as the total workforce is limited.

To achieve this reallocation gains, input sources such as labour and capital need to be reallocated at the lowest costs. The rigidities for firms to exit the market and for workers to switch jobs might hamper the process of adjustment to a higher aggregate productivity level.[42] Therefore, both labour market and capital market flexibility seem to be important policies for further success of policies affecting firm entry.

[42] It should be noted that it concerns exit costs, which are not accounted for by entrants. Generally, the size of the exit costs are known before entry, and the potential entrant accounts these with an expected probability to be forced to exit in its expected discounted profits after entry. In this case the exit costs have no additional impact as they have already been included in the entry decision. For empirical estimates of these costs, see Bartelsman et al., 2008.

Policies to improve labour market flexibility, as well as more streamlined bankruptcy procedures are policy options here to consider. However, the empirical evidence related to labour market regulation and productivity is scarce and provides no decisive answer. The effect of employment protection might have a positive impact on productivity as less flexibility of workers might enhance the incentives for employers to invest in specific qualities (see Deelen et al., 2005). Additionally, Acemoglu et al. (2002) add to this discussion the importance of distance to the frontier. Too much employment protection and discriminating outsiders might be bad for innovation if countries are close to the frontier. Moreover, we argue that equity issues like social cohesion and income distribution should be taken into account as well.

5.5 Are industry specific entry policies needed?

The theories discussed in chapter 4 operate with variables that may vary in size from industry to industry in practice. For example, *entry costs* in the pharmaceutical industry are larger than in the bakery industry, as the innovation costs for a new drug are larger than for a new bakery product (plus the economies of scale in R&D are likely to be greater in more traditional industries). Industries with a market leader can better *deter entry* than industries where market shares are more evenly distributed. For instance, entry deterrence is more likely in the soft drink industry where Coca Cola is the market leader than in the milk industry. The information asymmetry leading to a *capital market failure* is probably larger for an entrant in the new biotechnology industry than in the old mechanical machinery industry, as the risks to banks are higher when firms produce new, rather than existing, products. The *entry barriers for foreign enterprises* in the personal services industry are higher than in the car components industry, due to transportation costs and language barriers.

The question arises whether such industry heterogeneity should necessitate industry-specific policies or technology-specific policies with respect to firm entry. A purported case for such policies is based on the argument that quantitative effects of policy instruments may vary. Indeed, economic theories concerning entry consider a highly simplified economy or industry, omitting specific parameters. So, policy makers and competition authorities may choose to focus on specific industries or technologies where policy gains are largest.

This argument, however, does not necessarily imply the need for specific policies concerning entry. Above all, there must be a market failure for government to intervene the market. In the case of abuse of market power, the Dutch competition authority should take action to enforce fair competition in all Dutch sectors. Hence, this generic control function is most times case-specific in its implementation.

In the case of production externalities, the arguments for government intervention should be based on the evidence that the social rate of return is larger than the private rate of return. A priori, it is not reasonable to assume that the government knows better where new firms have to put their money in, where the production externalities are larger or which firm should enter the market (i.e. picking the winner). It is very difficult empirically to determine the size of market failures by industry, so that information requirements for specific policies may be overwhelming. Also, that entry will react differently to the same policy depending on the industry is irrelevant, because the reason for promoting entry is to allow actual entry to approach the optimal market outcome, which may well be industry-specific.

Finally, the theories reviewed in chapter 4 are unanimous on the effects of entry costs on productivity, which is the main objective of policies to promote firm entry. Therefore, these can be generic as long as they do not generate negative externalities. Market failures related to production externalities or network externalities (see box) that might be industry-specific. But, as we explained earlier, they can be targeted by policies other than promoting entry by reducing entry barriers.

Entry in network industries and network externalities

In network industries (such as in the telecommunications and software industries), consumer utility depends on the number of other consumers that make use of the same network called network effects. The existence of network effects does not automatically result in a market failure. A market failure called here network externalities emerges only if market parties are not able to internalise these effects. Network externalities result in market power if they give (incumbent) large suppliers a competitive advantage above new entrants or firms having significantly smaller networks. For the entrant, it will not be enough to have a product of better quality, or be more productive. The entrant must also be able to cover a large enough size of the network by its product. If a large number of consumers have already purchased the products offered by the incumbents, the entrant will find it especially difficult to enter the market and convince those consumers to switch to the new product. This problem can become even more acute if the incumbent's product is incompatible with that of the entrant.

Thus, in network industries entry can be very limited even in the absence of entry barriers. In order to stimulate entry, policy makers can enforce mandatory compatibility, open access, and open standards. These measures do have a positive effect on entry. Thus, we can argue that the existence of such an industry specific regulation does not alter the general policy conclusions in this section for other more 'traditional' industries.

5.6 Summary

This chapter derives implications for policies to promote productivity through firm entry. To do so, we apply a welfare economic approach assuming that the market outcomes are economically optimal from a welfare perspective as long as no market failures exist.

Policies that stimulate 'entrepreneurship' largely coincide with policies enhancing entry of (domestic) firms. The theoretical models showed that entry per se should not be the goal of economic policies, but affecting the size of entry costs can be a potential instrument to improve productivity through competition and/or innovation. The predicted direction of change followed by their implementation is seemingly easy: lower entry costs promote entry and in its turn entry leads to higher aggregate productivity through competition and/or innovation. The consequences for optimal policy design are, however, sometimes complex, not least because it requires that policy makers generally need much specific and detailed information. It is important to consider the pros and cons of policy interventions in cases where normal market outcomes can be mistaken for a market failure or much detailed information is required. In fact, an ex ante cost-benefit analysis needs to precede each intervention.

This chapter particularly considers three potential market failures: market power, capital market failures and production externalities related to knowledge spillovers and business stealing effects. Market power of incumbents occurs if incumbents can exercise market power to prevent entry or if institutional barriers protect them. A capital market failure is due to asymmetric information between the supplier of funds (e.g. banks) and potential entrants. R&D or innovation activities generate knowledge spillover effects creating benefits that cannot be fully appropriated by the inventor.

For each market failure we highlight the failure, the consequences, empirical evidence and policy options taking account of the policies already in place to repair them. In fact, the existing competition legislation should prevent the abuse of market power of incumbents. Preventing that abuse gives higher aggregate productivity and more welfare. This is, therefore, a no-regret policy. Well-designed intellectual property rights and innovation subsidies should address production externalities directly, not regulating entry. Those innovation instruments tackle the core problem of this market failure at its roots. Finally, capital rationing to new firms is often not a market failure. The supply of credit may be less due to new ventures being inherently more risky, which is a fair market outcome. Policy makers should carefully consider the information gap between the borrower and the potential entrant to assess the effectiveness of capital market policy (e.g., loan subsidies) in order to repair the market failure while acknowledging the uncertainty inherent in entry and its consequences for capital availability and costs.

Reducing (institutional) entry barriers seems to be the most promising channel for policy makers to improve productivity. In that respect, it is an interesting finding that the Netherlands has no top ranking on (components of) product market regulation and employment protection in an international perspective if countries are ranked from low to high entry barriers. This finding needs further consideration including the following issues. First, one should be careful to draw firm conclusions based on those rankings as rankings can be biased due to measurement issues. Second, the outcome can be the preferred outcome of a particular country given its choices with respect to efficiency and equity issues. Third, government interventions are only justified as long as social benefits are larger than the costs of intervention. Therefore, an ex ante cost-benefit analysis needs to precede each intervention. Such an analysis should also consider alternative entry policies that can affect productivity as not only entry costs determine whether a firm enters the market. Examples are policies that (ex ante) reduce exit costs and labour market rigidities related to market failures. A cost-benefit analysis is, however, beyond the scope of the document.

6 Concluding remarks

6.1 Conclusions

This document provides a review of recent theoretical and empirical literature on the relationship between entry, exit and productivity. The general conclusions from this review are as follows.

Empirical decomposition studies of productivity growth show that entry and exit of firms contribute to aggregate productivity growth considerably. The econometric evidence, however, provides an ambiguous correlation between business ownership and productivity growth. Moreover, this correlation does not say anything about causality.

Theory can say more about the causality between entry, exit and productivity. While entry, actual or the threat of, does encourage at least some of the incumbents to become more efficient, entry itself is driven by market opportunities arising either exogenously or as a result of entrants' own efforts (e.g., innovation activities). Theory clearly shows that high as well as low entry rates can go together with productivity growth. So, the message from theories to policy makers is that entry per se should not be the goal of economic policies, but affecting the size of entry costs can be a potential instrument to improve productivity through competition and/or innovation. Stated otherwise, it is productivity as an indicator of welfare that should be the main target for policy and not, for instance, the number of entrants, the intensity of competition or the number of innovations.

We have looked at policy options with respect to firm entry that might improve welfare through the repair of market failures, i.e. situations where markets do not deliver an efficient outcome (i.e. maximum welfare). We particularly considered three potential market failures: market power, capital market failures and production externalities related to knowledge spillovers and business stealing effects. Market power of incumbents occurs if incumbents can exercise market power to prevent entry or if institutional barriers protect them. A capital market failure is due to information asymmetry between the supplier of funds (e.g. banks) and potential entrants. R&D or innovation activities generate knowledge spillover effects creating benefits that cannot be fully appropriated by the inventor.

For each market failure we highlight the failure, the consequences, empirical evidence and policy options tacking account of the policies already in place to repair them. In fact, the existing competition legislation should prevent market power abuse of dominant firms. Well-designed intellectual property rights and innovation subsidies should address production externalities, not by regulating entry. Those instruments tackle the core problem of the market failure at its roots. Finally, capital rationing to new firms is often not a market failure, because the supply of credit may be less due to new ventures being inherently more risky, which is a fair market outcome.

Reducing (institutional) entry barriers seems to be the most promising channel for policy makers to improve productivity. In that respect, it is an interesting finding that the Netherlands has no top ranking on (components of) product market regulation and employment protection in an international perspective if countries are ranked from low to high entry barriers. This finding needs further consideration including the following issues. First, one should be careful to draw firm conclusions based on those rankings as rankings can be biased due to measurement issues. Second, the outcome can be the preferred outcome of a particular country given its choices with respect to efficiency and equity issues. Finally, government interventions are only justified as long as social benefits are larger than the costs of intervention. Therefore, an ex ante cost-benefit analysis needs to precede each intervention taking alternative options into account that can stimulate productivity through entry as well. That kind of analysis is beyond the scope of the document.

6.2 Questions for future research

The findings from this study suggest that reductions in institutional entry barriers might be an effective policy tool to stimulate productivity growth.[43] Still, several questions need to be addressed before introducing such policy. One of the main questions to be answered is whether this is the most effective policy instrument to improve productivity.

Robustness of evidence

As theory is ambiguous on the relationship between competition, innovation and productivity, the evidence rests on empirical studies. Most empirical evidence suggests that a higher threat of entry stimulates competition, and entry seems to encourage static efficiency and dynamic efficiency. With regard to the Netherlands, empirical evidence for an inverted-U link between competition and innovation is scarce. Hence, this relationship represents an interesting avenue for further empirical research based on the Dutch firm-level data.

Alternative entry options also need research

Alternative options for policy makers to stimulate productivity via entry are policies to reduce adjustment and exit costs, such as costs due to employment protection legislation. A priori, these costs are additive components in the decision of a firm to enter the market. However, the empirical evidence related to labour market regulation and productivity is also scarce and provides no decisive answer. For example, lower labour market regulation may also have a negative impact on productivity as more flexibility of workers might reduce the incentives for

[43] For your information, the CPB is already conducting research focused on the quantitative analysis of the macro-economic effects of policies aimed at education, innovation and research and development. This research also looks at how to build that into a growth model.

employers to invest in specific skills. Moreover, equity issues like social cohesion and income redistribution should be taken into account as well. Therefore, the interaction of labour market regulations and product market regulation and its impact on productivity is another interesting topic for further research.

That the Netherlands is doing fine in terms of productivity level in an international perspective, despite relatively low levels of entry and innovation expenditures, points to a range of other exogenous factors affecting productivity growth such as skills (e.g. educational and managerial skills) and knowledge spillovers. While the impact of many of those factors has been studied before, their interaction with firm entry (including the quality of the entrants) and its consequences for productivity makes an interesting topic for future research. So, the consequences of firm entry and exit for other aspects of business should be researched as well including flexibility within the firm.

References

Abernathy, W. and J. Utterback, 1978, Patterns of industrial innovation, *Technology review*, 80, pp. 41-47.

Acs, Z.J., D.B. Audretsch, P. Braunerhjelm and B. Carlsson, 2005, The missing link, The knowledge filter and entrepreneurship in endogenous growth, CEPR Discussion Paper 4783, CEPR.

Acemoglu, D., P. Aghion and F. Zilibotti, 2002, Distance to Frontier, Selection, and Economic Growth, NBER Working Paper 9066, The USA.

Aghion, P. and P. Howitt, 1992, A model of growth through creative destruction, *Econometrica*, vol. 60, nr. 2.

Aghion, P. and P. Howitt, 1997, A Schumpeterian perspective on growth and competition, Chapter 8, pp. 279-317, in Kreps D. and K. Wallis eds, *Advances in Economics, Theory and Applications*.

Aghion, P. and P. Howitt, 2006, Appropriate growth policy: a unifying framework, *Journal of the European Economic Association*, April-May 2006, 4 (2-3), pp. 269-314.

Aghion, P., R. Blundell, R. Griffith, P. Howitt and S. Prantl, 2004, Entry and productivity growth: Evidence from microlevel panel data, *Journal of the European Economic Association*, 2(2–3), pp. 265–276.

Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt, 2005a, Competition and Innovation: An Inverted-U Relationship, *Quarterly Journal of Economics,*120(2), pp. 701-728.

Aghion, P., R. Blundell, R. Griffith, P. Howitt and S. Prantl, 2005b, Firm Entry and Growth: Theory and Micro Evidence, *Journal of the European Economic Association*, 3(2-3), pp. 291-302.

Aghion, P., R. Blundell, R. Griffith, P. Howitt and S. Prantl, 2006, The effects of entry on incumbent innovation and productivity, NBER working paper 12027.

Audretsch, D.B., T. Aldridge and A. Oetl, 2006, The knowledge filter and economic growth: The role of scientist entrepreneurship, Discussion Papers on Entrepreneurship, Growth and Public Policy, Max Planck Institute of Economics.

Baily, M., C. Hulten and D. Campbell, 1992, Productivity dynamics in manufacturing plants. Brookings Papers on Economic Activity: Microeconomics 2, pp. 187-249.

Baldwin, J.R. and W. Gu, 2006, Competition, firm turnover and productivity growth, Statistics Canada Economic Analysis Research Paper 042.

Balk, B.M. and E. Hoogenboom-Spijker, 2003, The measurement and decomposition of productivity change: Exercises on the Netherlands' manufacturing industry, Statistics Netherlands, Discussion Paper 03001.

Barla, P. and S. Perelman, 1989, Technical efficiency in airlines under regulated and deregulated environments, *Annals of Public and Cooperative Economics*, 60 (1), pp. 103–124.

Bartelsman, E.J. and M. Doms, 2000, Understanding Productivity: Lessons from Longitudinal Microdata, *Journal of Economic Literature*, 38(3), pp. 569-594.

Bartelsman, E., S. Scarpetta and F. Schivardi, 2003, Comparative analysis of firm demographics and survival: Micro-level evidence for the OECD countries, OECD Economics working paper 348.

Bartelsman, E., J. Haltiwanger and S. Scarpettai, 2004, Microeconomic evidence of creative destruction in industrial and developing countries, The World Bank, Policy Research Working Paper 3464.

Bartelsman, E., E. Perotti and S. Scarpetta, 2008, Barriers to exit, experimentation and comparative advantage (forthcoming).

Bennett, M., P. de Bijl and M. Canoy, 2001, Future Policy in Telecommunications, An Analytical Framework, CPB Document 5.

Boot, A. and A. Schmeits, 2004, Imperfecties in de vermogensmarkt en overheidsbeleid, in: *Innovatie in Nederland, De markt draalt en de overheid faalt*, Preadviezen van de Koninklijke Vereniging voor de Stathuishoudkunde 2004, ISBN 90-807422-3-6.

Bunch, D. and R. Smiley, 1992, Who Deters Entry? Evidence on the Use of Strategic Entry Deterrents, *Review of Economics and Statistics*, 74(3), pp. 509-521.

Bushnell, J. and C. Wolfram, 2005, Ownership Change, Incentives and Plant Efficiency: The Divestiture of U.S. Electric Generation Plants, SCEM WP 140, US.

Cabral, L. 2002, *Introduction to industrial organization*, The MIT press, Cambridge Massachusetts, London, UK.

Carree, M.A. and A.R. Thurik, 2007, The lag structure of the impact of business ownership on economic performance in OECD countries, *Small Business Economics* (2008) 30, pp. 101-110.

Carree, M.A., van Stel, A., Thurik, R. and S. Wennekers, 2007, The relationship between economic development, and business ownership revisited, Tinbergen Institute discussion paper TI 2007-022/3.

Canton, E., B. Minne, A. Nieuwenhuis, B. Smid, and M. van der Steeg, 2005, Human capital, R&D, and competition in macroeconomic analysis, CPB Document 91.

Caves, R.E., 1998, Industrial Organization and New Findings on the Turnover and Mobility of Firms, *Journal of Economic Literature*, 36(4), pp. 1947-1982.

Chang, H.L. and F.F. Tang, 2001, An Empirical Study on Strategic Entry Barriers in Singapore, *Asia-Pacific Journal of Management*, 18(4), pp. 503-517.

Charnes A., W.W. Cooper and E. Rhodes, 1978, Measuring the efficiency of decision making units, *European Journal of Operations Research*, 2(6), pp. 429-444.

Coe, D.T. and E. Helpman, 1995, International R&D spillovers, *European Economic Review,* pp. 859-887.

Coe, D.T., E. Helpman and A.W. Hoffmaister 1997, North-South R&D spillovers, *Economic Journal,* pp. 134-149.

Coelli, T.J., D.S. Prasada Rao and G.E. Battese, 1998, *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Boston, the US.

Conway, P., V. Janod and G. Nicoletti, 2005, Product market regulation in OECD countries: 1998 to 2003, OECD economics department working paper 419.

Cornet, M. and B. Vroomen, 2005, Hoe effectief is extra fiscale stimulering van speur- en ontwikkelingswerk?, Effectmeting op basis van de natuurlijk-experimentmethode, CPB Document 103.

Cornet M., F. Huizinga, B. Minne and D. Webbink, 2006, Kansijk Kennisbeleid, CPB Document 124.

CPB, 2002, De pijlers onder de kenniseconomie; opties voor institutionele vernieuwing. The Hague, Special Publication 35.

Creusen, H., B. Vroomen, H.P. van der Wiel and A.E. Kuypers, 2006, Dutch retail trade on the rise? Relation between competition, innovation and productivity, CPB Document 137.

Deelen, A.P., E.L.W. Jongen and S. Visser, 2006, Employment protection legislation; lessons from theoretical and empirical studies for the Dutch case, CPB Document 135.

Disney, R., J. Haskel and Y. Heden, 2003, Restructuring and Productivity Growth in UK Manufacturing, *The Economic Journal*, 113 (July), pp. 666–694.

Driessen, G., M. Lijesen and M. Mulder, 2006, The impact of competition on productive efficiency in European railways, CPB Discussion paper 71.

Dunne T., M.J. Roberts and L. Samuelson, 1988, Patterns of Firm Entry and Exit in US Manufacturing Industries, *RAND Journal of Economics*, 19(4), pp. 495-515.

Dunne T., M.J. Roberts, L. and Samuelson, 1989, The Growth and Failure of US Manufacturing Plants, *Quarterly Journal of Economics*, 104(4), pp. 671-698.

EIM, 2007, WBSO Evaluation 2001-2005; Impacts, target group reach - and implementation.

Ericson, R. and A. Pakes, 1995, Markov-Perfect Industry Dynamics: A Framework for Empirical Work, *The Review of Economic Studies*, 62(1), pp. 53-82.

European Commission, 2003, Green Paper: Entrepreneurship in Europe.

European Commission, 2004, Enterprise - Action Plan: The European agenda for Entrepreneurship.

European Commission, 2005, Impact of market entry and exit on EU productivity and growth performance, European Economy Economic Papers, N 222, paper by Cincera, M. and O.Galgau.

European Commission, 2008, European Competitiveness report 2008.

Erken, H., P. Donselaar and R. Thurik, 2008, Total factor productivity and the role of entrepreneurship, Jena Economic Research Papers 2008-019.

Evans, D.S., 1987, The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries, *Journal of Industrial Economic*s, 35(4), 567-581.

Fabrizio, K., N. Rose and C. Wolfram, 2006, Does Competition Reduce Costs? Assessing the Impact of Regulatory Restructuring on U.S. Electric Generation Efficiency, working paper, http://faculty.haas.berkeley.edu/wolfram/Papers/frw.041306.pdf

Foster, L., J. Haltiwanger and C. Syverson, 2005, Relocation, firm turnover, and efficiency: Selection on productivity or profitability?, NBER Working Paper 11555.

Geroski, P.A., 1995, What Do We Know About Entry?, *International Journal of Industrial Organization*, 13, pp. 421-440.

Gort, M. and S. Klepper, 1982, Time Paths in the Diffusion of Product innovations, *Economic Journal*, 92(3), pp. 630-653.

Griliches, Z. and H. Regev, 1995, Firm productivity in Israeli industry, 1979-1988, *Journal of Econometrics*, 65, pp. 175-203.

Grossman G. and E. Helpman, 1991, *Innovation and growth in the global economy*, Cambridge MA, The MIT Press.

Hall, B. H., 2005, *The financing of innovation*, Blackwell Handbook of Technology, Oxford, Blackwell.

Haltiwanger, J., 1997, Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence, Federal Reserve Bank of St. Louis, *Economic Review*, 79, pp. 55-77.

Hellmann, T. and M. Puri, 2000, The Interaction between Product Market and Financing Strategy: The Role of Venture Capital, *Review of Financial Studies*, 13, pp. 959-584.

Hopenhayn, H.A., 1992, Entry, Exit, and Firm Dynamics in Long Run Equilibrium, *Econometrica*, 60(5), pp. 1127-1150.

Hopenhayn, H.A. and R. Rogerson., 1993, Job Turnover and Policy Evaluation: A General Equilibrium Analysis, *Journal of Political Economy*, 101(5), pp. 915-938.

Jacobs, B., R. Nahuis and P. J. G. Tang, 2002, Sectoral Productivity Growth and R&D Spillovers in the Netherlands, *De Economist*, 150, pp. 181-210.

Jamsab, T. and M. Pollitt, 2001, Benchmarking and Regulation: International Electricity Experience, *Utilities Policy*, 9, pp. 107-130.

Jovanovic, B., 1982, Selection and the Evolution of Industry, *Econometrica*, 50(3), 649-670.

Klepper, S., 1996, Entry, Exit, Growth and Innovation over the Product Life Cycle, American *Economic Review*, 86(3), pp. 562-583.

Laffont, J.J. and J. Tirole, 1993, *A Theory of Incentives in Procurement and Regulation*, MIT Press, Princeton.

Lerner, J., 1996, The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program, NBER Working Paper 5753.

Mankiw, N.G. and M. Whinston, 1986, Free Entry and Social Inefficiency, *Rand Journal of Economics* 17, Spring 1986, pp. 48-58.

Ministry of Economic Affairs, 2003, Final progress report MDW, June 2003, ME/B4 3020265.

Nicoletti, G. and S. Scarpetta, 2003. Regulation, productivity and growth: OECD evidence, OECD Working Paper 347.

OECD, 2001, OECD Economic Outlook 69, Chapter 7: Productivity and firm dynamics: Evidence from microdata.

OECD, 2008, Measuring entrepreneurship: a digest of indicators, OECD-Eurostat Entrepreneurship Indicators Program.

Olley, G.S. and A. Pakes, 1996, The Dynamics of Productivity in the Telecommunications Equipment Industry, *Econometrica*, 64(6), pp. 1263-1297.

Pakes, A. and R. Ericson, 1998, Empirical Implications of Alternative Models of Firm Dynamics, *Journal of Economic Theory*, 79(1), pp. 1-45.

Pakes, A. and P. McGuire, 1994, Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model, *RAND Journal of Economics*, 25(4), pp. 555-589.

Park, J., 2004, International student flows and R&D spillovers, *Economic Letters*, volume 82, issue 3, March 2004, pp. 315-320.

Santarelli E. and M. Vivarelli, 2006, Entrepreneurship and the process of firms' entry, survival and growth, IZA Discussion Paper 2475, December.

Steiner, F., 2001, Regulation, Industry Structure and Performance in the Electricity Supply Industry, OECD Economics Department Working Papers 238, OECD Publishing.

Statistics Netherlands, 2008, Het Nederlandse ondernemingsklimaat in cijfers 2008.

Stiglitz, J. E., 1968, A Note on Technical Choice under Full Employment in a Socialist Economy, *Economic Journal*, 78, pp. 603-609.

Tybout, J.R., 2000, Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?, *Journal of Economic Literature*, 38(1), pp.11-44.

Utterback, J. and W. Abernathy, 1975, A dynamic model of process and product innovation, *Omeg*a, 3(6), pp. 639-656.

Verheul, I., L.M. Uhlaner and A.R. Thurik, 2002, Gender, entrepreneurial activity and entrepreneurial self-image: new insight into the question: what is an entrepreneur? In *Proceedings of the 2002 Small Business and Entrepreneurship Development Conference* (pp. 452-463).

Wiel, H.P. van der, 1999, Firm turnover in Dutch business services: The effect on labour productvity, CPB Research Memorandum 159.

Wiel, H.P. van der, H. Creusen, G. van Leeuwen and E van der Pijll, 2008, Cross your border and look around, CPB Document 170.

Appendix A: Decompositions of productivity growth to account for firm entry and exit

Suppose firm i produces in period t the amount of output Y_{it} using L_{it} units of labour. Then the average productivity in the industry is

$$P_t = \frac{\sum_{i=1}^{N} Y_{it}}{\sum_{i=1}^{N} L_{it}} = \frac{\sum \frac{Y_{it}}{L_{it}} \cdot L_{it}}{\sum L_{it}} = \sum p_{it} s_{it}, \tag{A.1}$$

where p_{it} is firm i's labour productivity in period t and s_{it} is its share in total employment. The average productivity change between periods t and $t-1$ is

$$\triangle P_t = P_t - P_{t-1} = \sum p_{it} s_{it} - \sum p_{it-1} s_{it-1}. \tag{A.2}$$

The population of firms in periods t and $t-1$ consists of continuing firms (C), present in both t and $t-1$, firms entering in period t (E) and those leaving the market in $t-1$ (X). We can then rewrite equation (2) as

$$\triangle P_t = \sum_{i \in E} p_{it} s_{it} + \sum_{i \in C} s_{it-1}(p_{it} - p_{it-1}) + \sum_{i \in C} (s_{it} - s_{it-1}) p_{it} - \sum_{i \in X} s_{it-1} p_{it-1} \tag{A.3}$$

This decomposition is due to Baily et al. (1992) and is derived from equation (2) by a simple rearranging of terms. Here the first term is the contribution of entering firms, the second term is the contribution of within-firm productivity growth, the third term is due to changes in employment shares, and the fourth is due to firms exiting the market.

Another decomposition was proposed in Haltiwanger (1997):

$$\begin{aligned} \triangle P_t = {} & \sum_{i \in E} s_{it}(p_{it} - P_{t-1}) + \sum_{i \in C} s_{it-1}(p_{it} - p_{it-1}) + \sum_{i \in C} (s_{it} - s_{it-1})(p_{it-1} - P_{t-1}) \\ & + \sum_{i \in C} (s_{it} - s_{it-1})(p_{it} - p_{it-1}) - \sum_{i \in X} s_{it-1}(p_{it-1} - P_{t-1}) \end{aligned} \tag{A.4}$$

Again, (4) is a rewrite of (2), but in a way different from Baily et al. (1992). The first and last terms of Haltiwanger's decomposition in (4) represent the contribution of firm entry and exit, positive if entering firms are more productive and exiting firms are less productive than industry average in the last period. The second term is the familiar within-firm productivity growth. The third term defines the contribution of employment reallocation, positive if firms more productive than average in the last period gain a large share in total employment in the current period. The fourth term is 'covariance', positive if firms showing positive productivity growth gain larger share in total employment.

Another popular decomposition is due to Griliches and Regev (1995):

$$\begin{aligned}\triangle P_t &= \sum_{i \in E} s_{it}\left(p_{it} - \frac{P_t + P_{t-1}}{2}\right) + \sum_{i \in C} \frac{s_{it} + s_{it-1}}{2}(p_{it} - p_{it-1}) \\ &+ \sum_{i \in C}(s_{it} - s_{it-1})\left(\frac{p_{it} + p_{it-1}}{2} - \frac{P_t + P_{t-1}}{2}\right) - \sum_{i \in X} s_{it-1}\left(p_{it-1} - \frac{P_t + P_{t-1}}{2}\right)\end{aligned} \qquad \text{(A.5)}$$

(5) resembles (4), save that the reference point now is the two-period average, not last period's, labour productivity. Here too, entry and exit contribute to productivity growth positively if the entering firms are more productive, and the exiting firms are less productive, than the industry average. The contributions of within-firm productivity growth and employment reallocation between firms are defined similarly. Averaging labour productivity and employment shares over two consecutive periods saves the need to have the covariance term. As before, (5) can be reduced to the original productivity growth equation (2).

If all the decompositions are just re-writes of the productivity growth identity (2) which one should we use? Recall that the contribution of firm entry and exit to productivity growth is calculated in different decompositions against different benchmarks. Thus, in (3) the contributions of firm entry and exit are taken as such, whereas in (4) there is a benchmark, P_{t-1}, and in (5) it is $(P_t + P_{t-1})/2$. This benchmark productivity is meant to be an estimate of industry average productivity in the absence of entry or exit, subject to certain assumptions. It is sometimes called counterfactual productivity (e.g., Baldwin and Gu 2006). Examining the assumptions underlying counterfactual productivity is important for choosing the 'right' decomposition.

The counterfactuals for all the decompositions above assume constancy of the employment shares in the absence of entry or exit, and that the continuing firms' productivity stays at its actual level and the exiting firms' at some average level. However, because employment shares stay constant while labour productivity between the continuing and exiting firms differs, all the counterfactuals reviewed so far imply changes in output shares in period t compared with $t-1$, thus not fully controlling for the effect of competition among the incumbents on average productivity. As a result, the contribution of the within-firm productivity growth may be overestimated, and that of entry and exit underestimated. Baldwin and Gu (2006) address the problem of implicit changes in output shares of the incumbents in the counterfactual by holding output shares in t and $t-1$ constant. The productivity growth decomposition consistent with their counterfactual is

$$\triangle P_t = \sum_{i\in C}(s_{it}-\hat{s}_{it})(p_{it}-p_{xt-1})+\sum_{i\in E}s_{it}(p_{it}-p_{xt-1}) \\ +\sum_{i\in C}\frac{1}{2}(\hat{s}_{it}+s_{it-1})(p_{it}-p_{it-1})+\sum_{i\in C}\left(\hat{s}_{it-1}-s_{it-1}\right)\left(\frac{1}{2}(p_{it}+p_{it-1})-p_{xt-1}\right) \tag{A.6}$$

Here we have a benchmark, p_{xt-1}, which is unique, unlike in the previous decompositions, and a counterfactual share of employment consistent with constant output share, $\hat{s}$. The first term in (6) represents the contribution to the productivity growth of output share changes. It is positive if firms more productive than an average exiting firm acquire larger market shares. The second term is the total contribution of entry and exit, positive if new entrants are on average more productive than the exiting firms. Note that there are no separate contributions of entry and exit – the market share of the exiting firms is taken either by the continuing firms, for which the first term has already accounted, or by the new entrants (the second term). The third term in (6) is the contribution of the within-firm productivity growth to the total. The fourth is the 'covariance' term. It is approaching zero when all the continuing firms show the same productivity growth (for reasonable, not explosive growth rates), negative when firms with less than average productivity gain a larger share in total employment, and positive otherwise.