4 ORGANIZATIONAL ECOLOGY AND INDUSTRIAL ORGANIZATIONS; FRAMEWORKS FOR PATH DEPENDENCY AND LOCK-IN OF OLD FIRMS

In Chapter 2, the firm was defined, both from its relationship with its environment and as a unit of analysis. As explained in that chapter, the latter vision fits into the theory of organisational ecology (Van Wissen, 2002), which Hannan and Freeman (1989) introduced with a focus on the boundaries between the organisation (firm) and the market. From their theory of organisational ecology, which is the basis of the demography of firms, emerged the necessity to view the firm as a unit of analysis. Nevertheless, the interactions that a firm has with its environment are still important (Carroll and Hannan, 2000).

The insights from organizational ecology (Hannan and Freeman, 1989) have already been briefly introduced in Chapter 2. Other pioneers in this field of study are Nelson and Winter (1982), who advocate the evolutional approach to economic development through the theory of industrial organization. The theory of organizational ecology was derived from sociology, and takes the study of entry and exit data as the main theme. The number of firms (density) in an industrial sector is the core variable in the explanation of entry and exit rates of firms. The relationship between these two concepts is described in the so-called 'density model of competition and legitimation'. The potential limit to the number of firms in a certain sector is called the carrying capacity. One of the shortcomings of organizational ecology is the lack of attention to individual firms. The school of industrial organization does pay attention to individual firms; however, in the literature on industrial organization, too little attention is given to the relationship between industry dynamics (entry and exit) and the age of firms. These concepts will be discussed in more detail in the following sections. Both theories have the firm itself as the unit of analysis, and also incorporate the interactions with its surrounding environment (Swaminathan, 1996). The evolutionary character of these theories also makes use of the concepts of selection, adaptation, inertia and path dependence (Antonelli, 1997; Forday, 1997).

This chapter will discuss organizational ecology and, in less detail, also industrial organizations and then apply the theoretical insights, such as the 'density model of competition and legitimation', to a conceptual framework for explaining the long-term survival of firms incorporating the use of selection, lock-in and inertia. Further, attention is given to the evolutionary notion of path dependence. All the described concepts will be viewed from a spatial angle and developed into testable hypotheses in order to investigate some of the characteristics of (old) firms in terms of the questions posed in Chapter 1. As Krugman (1991a, pp.80) argues: "if there is

a single area of economics in which path dependence is unmistakable, it is in economic geography – the location of production in space. The long shadow cast by history over location is apparent in all scales, from the smallest to the largest".

4.1 Introduction

Two theories that provide insights into the entry and exit of firms are the industrial organization and organizational ecology ones. Nelson and Winter (1988) introduce a broad evolutionary perspective that can be used to analyse the various aspects of economic change, such as the response of firms and industries to changed market conditions. The theory of industrial organization is concerned with the environmental settings in which firms operate, and how firms behave in these settings: as producers, as sellers and as buyers. In other words, there is no internal consideration: firms are viewed in terms of their market behaviour. A major emphasis is placed on the competitive forces in the various industries and markets. Correspondingly, the analysis takes place in the industry, or in the competing group of firms, rather than in the individual firm (Davies and Lyons, 1988). This focus on evolutionary change provides a dynamic explanation of variety and diversity (Peneder, 2001).
Organizational ecology focuses on organizations (firms) in trying to explain long-term social evolution, particularly the rise and fall of organizational populations. It views the market as a field for organizational actions. "The organization population is defined broadly as the set of organizations characterized by a particular organizational form and dependent on a common set of material and social resources" (Carroll, 1997, pp.2). This way of thinking encompasses a very specific research design that involves going back to the origins of a specific organizational population or industry, and describing all the organizations that enter or exit from it using a prescribed definition of organizational form (Carroll, 1997). The core variable in organizational ecology is the number of firms (density) in an industrial sector since, it is argued, this explains entry and exit rates. The relationship between these two concepts is described in the so-called 'density model of competition and legitimation'. The potential limit to the number of firms in a certain sector is called the carrying capacity (Hannan and Carroll, 1995). One of the main shortcomings of organizational ecology is its lack of attention to individual firms. The school of industrial organization does pay attention to individual firms but, in the literature on industrial organization, too little attention is given to the relationship between industry dynamics (the frequency of entry/exit) and the age of firms. In Section 4.1, firstly, an evolutionary view on the development and survival of firms is given from the industrial organization and organizational ecology perspective. Secondly, the density model of competition and legitimation is described in more detail, followed by the liabilities that can arise from this for firms of different ages. Section 4.1 then concludes with a spatial description of the earlier described

concepts in the context of old firms. In Section 4.2, the evolutionary concept of path dependence is described. First a general overview, followed by the incorporation of this concept in organizational ecology and industrial organization theories. Then, a discussion on age, inertia and lock-in follows: are these concepts path-dependent? Section 4.2 concludes with a discussion on spatial path dependence. In Section 4.3, the spatial implications of this evolutionary view of old firms are discussed. The variables influencing long-term firm survival are further elaborated, and the section ends with several hypotheses, for later testing, about old firm survival.

4.2 An evolutionary view on development and survival of firms

In this section, a general overview of the two introduced theories is first given, followed by an explanation of the model of density dependence, and the diverse liabilities that indicate survival chances of firms at different ages. The section ends with an elaboration of three concepts: adaptation, selection and structural inertia. The interactions among these three concepts are described along with the way these concepts might explain long-term firm survival.

4.2.1 Industrial organization theory

Industrial organization theory emphasizes entry and exit (the birth and death of firms), variations in size, and market share (mobility) and changes in the control of firms, as demonstrated by a firm's behaviour in the market. The main focus is not on the individual firm, but on the total industry or group of interacting individual firms (Caves, 1998). Industrial organization studies contribute to the understanding of how a market works from birth to maturity, and is associated with firms' behaviours. Recent studies have enhanced the insights into industrial dynamics. In general, the results show considerable short-term entry and exit turnover in most industrial markets; while, in the longer term, the market structure is relatively stable. Further, firms may behave in a competitive way because of the danger posed by potential entries. In other words, not only the number of existing firms, but also the potential number of new firms, determines the competition forces (Van Kranenburg, 1999).

Alongside competition, both growth and size are important determinants of the mortality and survival of firms. Following a study on mobility, entry and exit, Caves (1998) came to the following conclusions. Mean growth rates of surviving firms are not independent of their size, but tend to decline with size and with the unit's age for a given size. The entry and exit of firms are closely linked with growth-size relationships. Entry is more likely to occur in small size classes, and the likelihood of a unit's exit declines with size. Growth rates are almost independent of size among larger firms. However, the failure to sort firms into

industry groups deprives the results of clear implications. The evidence on entrants' growth and failure rates suggests the processes by which firms can succeed and that they do not initially position themselves with the unique, optimal size. Industries with high entry rates will also show high exit rates. The survival of mature firms is determined by the extent of depreciation or obsolescence in their original cumulative learning capabilities (Caves, 1998). When a firm ages and grows, it becomes more certain about its costs and, therefore, the mean and variance of its growth rate should decline. Although many industrial organization studies have contributed considerably to the understanding of the various market structures and firm behaviours, they are not sufficient to explain the evolution of industries. The industrial organization literature, it is argued, pays too little attention to the relationship between industry dynamics and the age of the industry (Van Kranenburg, 1999).

4.2.2 Organizational ecology theory

The development of organizational ecology theory took place in the late 1970s (Amburgey and Rao, 1996). Studying the ecology of organizations (and firms) is an approach that builds upon general ecological and evolutionary models of change in populations of organizations (firms). The goal of this perspective is to understand the forces that shape the structures of organizations over long period of time (Hannan and Freeman, 1989). Organizational ecology treats organizations as complex systems with strong limitations on flexibility and speed of response. The features of the theory are as follows. First, ecological models are used as frameworks to study these sociological processes. Second, the theory concentrates on dynamics, particularly on the dynamics of processes that shape entry and exit rates in the populations of organizations, and these processes are studied over long periods of time. Third, in the process of studying such dynamics, the theory considers entire populations of organizations over the full history of these populations (Hannan and Freeman, 1989). This is in line with the concept of path dependence as will be described in Section 4.3 (Antonelli, 1997)
Reasons for the development of the theory of organizational ecology were an interest in the causes of vital events (birth, death, and mobility of firms and organizations), as well as a desire to structure the connections between organization theory and social history. Research at the population level leads to an interest in history since the study of population dynamics frequently requires an analysis over long periods of time. Furthermore, it leads to an understanding of the institutional context of organizational populations (Hannan and Freeman, 1989). According to Carroll (1988), the existing organizational theories were ignorant of the founding and mortality processes. The focus should be more on the vital rates, the entries and exits, which are the distinguishing features of ecology, and which could be found in the present theory. Singh (1990, pp.11) defines the central theme

of the theory of organizational ecology as: "the investigation of how social environments shape rates of creation and death of organizational forms, rates of organizational founding and mortality, and rates of change in organizational forms." Organizational ecology does not subscribe to the adaptation model of organizational change: it challenges many of the fundamental features of the adaptation model and offers instead a selection model of organizational change. In the approach of the selection model, adaptive change by firms is not impossible but is seriously constrained. In this view, strong, inertial forces characterize organizations, and these forces are seen to limit the flexibility to change. From this viewpoint, then, most organizational changes occurring in any historical period are the result of processes of organizational selection and replacement, rather than internal firm changes and adaptations. In this way, organizational ecology forces attention towards patterns of organizational founding, mortality and survival: the visible outcomes of the selection processes (Carroll, 1988).
The most important element in organizational ecology is the population of organizations. The crucial assumption is that organizational populations can be defined so that they have a unitary character: the members of the population must have a mutual understanding with respect to the processes of interest (creation, dissolution and transformation) (Amburgey and Rao, 1996). However, empirical studies have found similar outcomes for different types of industrial groups, implying that the general assumption might well be applied to a more heterogeneous population (Amburgey and Rao, 1996). A criticism of the theory of organizational ecology is that, in the population set, very large and very small firms within one industry are seen as one homogeneous population of firms, despite the fact that very large firms have different routines and characteristics than small firms and, through this, have very different effects within the population (Winter, 1990). The reason for defining the population in this way is that the new and small firms are potentially competitive with the large firms. Furthermore, the large firms in the population started as small firms. Excluding the small firms from the population until they become competitively significant would introduce selection bias and obscure the many underlying processes of evolution, whatever the eventual outcome (Winter, 1990).

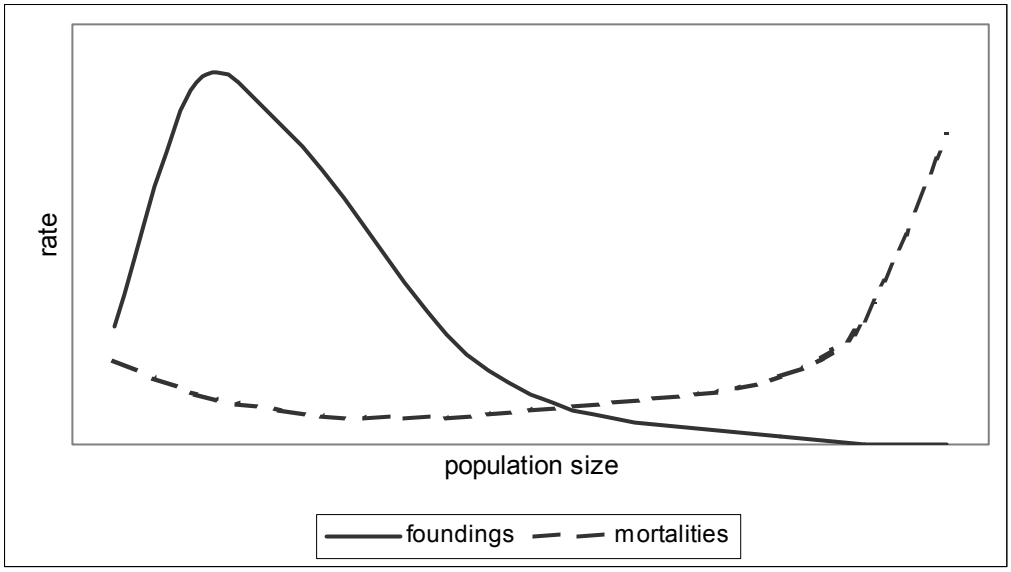
4.2.3 Density model of competition and legitimation

The density model of competition and legitimation identifies a relationship between entry and exit rates. The number of firms (density) in an industrial sector is the core variable in the explanation in the entry and exit rates (Hannan and Carroll, 1995). In this model, insights from both the organizational ecology theory and the industrial organization theory are used. The model assumes that a change in populations occurs mainly through the selective replacement of different organizations, rather than through the adaptations of individual organizations.

Carroll (1997) assumes that there are two forces in selection: social legitimation and diffuse competition. Social legitimation refers to the taken-for-granted nature of a particular firm as the 'natural' way of doing certain things. Diffuse competition refers to the competition that arises when organizations depend on the same set of finite resources (for example skilled labour). Diffuse competition implies that the viability of particular types of organizations will fall when many firms acquire these resources (Carrol, 1997). Low or absent legitimation means that managing will become difficult: capital suppliers are cautious; suppliers and customers need to be informed; employees may be hard to find and hire; and often institutional rules have to be changed. As more firms appear, legitimation increases. Initially, when the number of firms is low, the effects on legitimation of adding another firm is large. When many firms are present in the population, legitimation increases only little, or not at all, when the density rises. To put it succinctly, the legitimation of an organizational population increases with density at a decreasing rate (Carroll, 1997; Carroll and Hannan, 2000). This interaction between competition and legitimation in the model can be observed in the births and deaths of firms in a particular population. When legitimation increases, entrepreneurs see an opportunity and are more willing to start such a type of business. Firms founded during periods of growing legitimation also find it easier to attract capital, suppliers, customers and employees. These firms encounter relatively few institutional obstacles. With the rise of legitimation, not only do founding rates increase but mortality rates also decline. By contrast, if diffuse competition in an organizational population intensifies quickly, entrepreneurs become cautious (Carroll, 1997). This interaction process is illustrated be seen in Figure 4.1.

When density is low, competition is minimal. Competition intensifies as the size of the population increases, resulting in a higher likelihood of mortality for some of the population members. Decreasing birth rates and/or increasing death rates are the result, slowly causing the population's overall growth rate to decline and eventually fall to zero. A continuing increase in the number of deaths indicated an unhealthy environment for firms. The effect of density on the mortality rate is predicted to be negative at low densities and positive at high densities (Hannan and Carroll, 1995). Entrepreneurs respond to an unhealthy environment by not starting up new firms, and hence the number of foundings decline. Similarly, an initial increase in the creation of new firms encourages the creation of other firms, since this is a sign of a supportive environment (Aldrich and Staber, 1988). This competition can also be indirect; the growth of one population of organizations may reduce the growth of another population even though the members of these two populations never interact directly (Hannan and Freeman, 1989).

Figure 4.1: Founding and mortality rates according to the density dependence model



Source: Carroll (1997, pp.128)

Firms founded in a period of high density are expected to have higher age-specific mortality rates. Resources can be scarce in high-density periods, conditions that could cause mortality in weaker organizations (Swaminathan, 1996). Furthermore, Bigelow et al. (1997) found that the effects are age-dependent: the effects of density diminish with age. They found empirical results for this in very different sets of industries and the outcomes were similar over long time periods. The density dependence model does predict similar patterns of entry and exit for different industries and perhaps, therefore, the prediction can also be applied to a more heterogeneous group of firms (compare Bigalow et al, 1997) since markets also can be defined independently of the particular firms that inhabit them (Geroksi, 2001). From the results in Chapter 3, it is clear that the density dependence model is no longer effective after a certain period. The carrying capacity is reached and entry and exit rates become equal. The old firms in this investigation are the outcome of this cycle, these old firms have survived so long that they passed through the entire development cycle as shown in Figure 4.1. The survivors can be seen as the outcome of the density dependence model.

Van Wissen (2004) compares this model of density dependence with the concept of agglomeration economies, since both involve a positive feedback between size of the population and growth. However, van Wissen emphasizes that in the density dependence model spatial heterogeneity does not exist, and this is a serious shortcoming since firms do not only differ in size and activity, but also in geographical distribution. Other criticisms of the density model are: first, the concept of social legitimation is not very clear, it is hard to find direct measures for

this concept and it should be broadened to include legal and sociopolitical legitimacy (Baum, 1995). Second, the density variable is an incomplete way to measure evolution. The variable should also include, for example, size or location (Van Wissen, 2004). Third, the diffusion competition component could be better replaced by a direct competition component that would give a better representation of evolution (Carroll, 1997). One of the advantages of the density model is its generalizability since it can be applied to very diverse populations of firms, including firms of all sizes (Winter, 1990; Amburgey and Rao, 1996).

4.2.4 Liabilities and aging firms

In population ecology several rules-of-thumb are to be found, that are called liabilities. Firms have a lower or higher likelihood of survival at specific times during their lives. When firms are young, the 'the liability of newness' hypothesis argues that new firms in a population have a lack of social agreement, lack stability and insufficient resources. These limitations increase the risk of failure for new firms. Empirical evidence indicates that firm mortality tends to decline with increased size, and therefore large firms are expected to be less vulnerable to the risk of failure. Older firms are generally also the larger firms (Baum and Oliver, 1991). Furthermore, the liability of newness argument links problems faced by young organizations, such as barriers to entry, inability to attract skilled labour, and difficulties wih internal organization, to an increased risk of failure. Hannan and Freeman (1989) add that new firms also fail at higher rates than old ones, because they are more vulnerable being strangers in the population and having to create roles and routines: that is legitimation. Therefore, these new firms cannot rely on the generalized skills and experience available in the labour market. Hence mortality rates will drop sharply with age, even when the initial size of a firm is taken into account. The possibility exists that age variation in mortality rates is deceptive in the sense that it only reflects the operation of unobserved but age-dependent heterogeneity in cohorts of firms.
Next to the liability of newness, liabilities of adolescence and aging (senescence & obsolescence) also exist. The liability of adolescence argues that when firms have survived the liability of newness, they seldom fail until the initial resources become scarce or the initial investments need to be replaced. In this situation, the failure rates in the population will increase: firms need to realise in time that they need to adapt to changing circumstances; unsuccessful firms will fail and those firms that undergo a learning process and evaluate their performance wil survive (Brüderl and Schüssler, 1990). This brings the 'adaptive behaviour' argument back into the discussion. Firms that adapt have increased survival chances (Thornhill and Amit, 2003). Ranger-Moore (1997) identifies another liability, the so-called liability of aging, which can be divided into two types: senescence and obsolescence. Both liabilities occur in firms that 'grow old' and have survived the first two liabilities.

According to Ranger-Moore, old and large firms suffer from time-dependent processes and become less efficient and less effective, and loose their ability to respond adequately to changing conditions. Firms that experience senescence are, in other words, just growing older, and their internal processes age and lead to performance loss, just as in biological organisms. Those firms that experience obsolescence have become outdated or antiquated. In obsolescent firms, the internal processes become inert and are less able to respond to changing conditions. There are increased likelihoods of mismatches with resources, capabilities and the demands of competitive environment (Thornhill and Amit, 2003). The notion of inertia in older firms will be further elaborated in Section 4.1.5. Amburgey et al. (1993) use the same idea, but describe it differently. In older firms, internal roles and formal structures become more established and standardized over time. These old firms have lived so long that they have developed routines for internal and external linkages. So if changes in the environment disrupt the firm, the internal and external patterns of the old firm will suffer heavily. Furthermore, Hannan and Freeman (1984 and 1989) add that every attempt at organizational change may end in a so-called 're-setting of the liability of newness-clock', and so by adapting and changing, the firm has to start over every time and renew its risks of failure and termination once more. This is an argument against adaptation favouring survival. With these liabilities, the age of a firm can be seen as a proxy for the internal organizational processes in the firm that evolve over time (Thornhill and Amit, 2003).

The mortality rates of firms in a particular cohort may depend not only on the age of the firm, as suggested by the diverse liabilities, but also on the interaction between age and initial conditions, such as the environmental ones, at founding (Swaminathan, 1996). This is a first step towards the idea of path dependence (Liebowitz and Margolis, 1996). This is also strongly dependent on the initial conditions as will be discussed later. Empirical testing of the liabilities of aging by Ranger-Moore (1997) indicated that the relative effects of age are greater in obsolescence than in senescence. This implies that inert and older firms will only be in danger of exiting in turbulent times. During calm periods these firms have the advantages of reliability and accountability. Furthermore, a firm needs inertia to survive, to gain sufficient levels of reliability and accountability. However, if the levels of inertia are too high, a firm cannot keep up with changing conditions (Ranger-Moore, 1997).

4.2.5 Adaptation, selection and structural inertia in old firms

From the previous sections it is clear that some concepts are seen as important in the discussion about the survival of old firms from an evolutionary perspective. These are adaptation, or selection and structural inertia. These concepts are not mutually exclusive, but are to an extent accumulative and can occur at the same

time within the same firm (Amburgey, et al., 1993). With the aging of firms, these processes grow stronger within a firm and are influenced by history. No firm exists for a very long time without ever having adapted. The development of routines and habits by a firm results in inert behaviour and this must result in a positive selection by the environment since these old firms are the survivors. The main issue for these concepts is the level at which one looks at the firm: either from the population level - selectivity; or from the individual firm responding to its environment - adaptation and inertia. The latter also gives an opening to the study of a more heterogeneous group of firms. With the incorporation of history in this evolutionary view, not only is the survival of firms studied, but also the description of the experience that firms go through as they inevitably change over time (Geroski, 2001). According to Bruderer and Singh (1996), environmental selection influences adaptation; and adaptation, in turn, influences the selection process. Learning and selection are fundamentally interdependent processes because adaptation enhances inertia, just as inertia accelerates the process of environmental selection. There are three processes that influence organizational evolution: 1) the occurrence of variation, or the birth of innovative organizational forms; 2) firms learning as an example of adaptation and the maintenance of a positively selected option; 3) selection, that is the death of firms related to their fitness or consistent criteria for selection (Aldrich and Pfeffer, 1976; Bruderer and Singh, 1996). These three processes all occur simultaneously in the population of firms (Freeman, 1981). This selection process can be called evolution, but is also titled natural or ecological selection. Aldrich and Pfeffer (1976, pp.80) prefer natural selection to evolution, because evolution implies that there is a movement towards higher forms of social organization, or to better firms. However, this is not the case: "The process of natural selection moves towards a better fit of the firms with the environment, nothing more." Alongside the model of natural selection, Aldrich and Pfeffer (1982) describe the model of political economy, that argues for greater attention to a firm's internal political decision-making processes and to the perspective that firms seek in managing or strategically adapting to their environment. Both models agree on the importance of a firm's environment for understanding its decisions and structures (Aldrich and Pfeffer, 1976). However, when considering the long-term time frame this latter model is omitted.

According to the model of natural selection, selection is based on the survival of different structural forms, rather than on the adaptation of a single form. Furthermore, the selection process does distinguish between the survival of long-existing structures and of newly emerged firms. This is because the length of a firm's 'generation' is not known, if such a thing exists. In the long run, only those firms that best fit the environment will survive (Aldrich and Pfeffer, 1976). Hence, in investigating a group of old firms, one can expect to see common characteristics in structural form, since all these firms have been positively selected. This fitness of the firms to the environment reflects, according to Freeman (1981), how many

firms of a given kind can be observed in a locally bounded area. He described two characteristics that influence this fitness. First, the way that a firm acquires resources and, second, how many similar firms are in the same bounded area (competition). Aldrich and Pfeffer (1976) add that, for the survival of firms, economies of scale, cost barriers and product differentiation are also important. Pfeffer (1982), furthermore, argues that the market of buyers and suppliers have an influence on the survival of firms. This raises the issue of legitimation again. As legitimation rises, not only do organizational founding rates increase but also mortality declines (Carroll, 1997), compare Section 4.1.3. Institutional theorists argue that a firm is more likely to survive if it obtains legitimacy and social support.

Much research has been done on age dependency in firm survival; especially on the short-term time frame, and many populations exhibit a liability of newness. Stinchcombe already in 1965 wrote that, in any population, new organisations are more likely to die than older organisations and, at any age, organisations of a new form are more likely to die than organisations of an old form (Freeman, 1990). At the same time, Swaminathan (1996) argues that firms who survive the first years of their existence in a 'hard' environment will have higher survival-rates in the long term, they have been 'hardened by battle'. Jovanovic's model (1982) predicts that firm survival will increase with age and size of a firm. However, Agarwal (1997) argues that this model is unrealistic over an extended span of time because it does not allow for a mutation of the environment in which the firm operates and competes. Also Boschma and Lambooy (1999) argue for such a 'context-dependency' for firms and other organisations. They claim that regions are regarded as rather stable homogenous entities in terms of their collective knowledge, institutional structures and social conventions. Amburgey et al. (1993) call this the historical and structural embeddedness of relationships. Within this embeddedness, firms are resistant to change, and these relationships can turn into long-term dependencies that constrain the behaviour of a firm. Romo and Schwartz (1995), find this embeddedness in the migration behaviour of firms. They find that embedded firms do not leave the original region, even when cost structures in other regions of the country or the world might be more beneficial. Whichever, the survival of firms has to be studied from two angles. First, the individual firm and its inert and/or adaptive behaviour and second, the firm in its population, the selection process of the market and the environment.

Selectivity in entry and exit from a population, in combination with internal change of the surviving members, is a relevant field of study in firm populations (Van Wissen, 2002). However, the influence of the environment and the influence of other members of the populations do not fully explain why some firms become so much older than others. The effect of organisational age on demographic events remains to be fully understood, and also needs to be studied on the level of the individual firm. On the one hand, ageing implies learning, and thus becoming

better equipped. On the other hand, ageing may mean increased structural inertia, leading perhaps to obsolescence (Van Wissen, 2002; Brüderl and Schüssler, 1990; Ranger-Moore, 1997). Inertia, or resistance to change, is one of the main forces determining firm behaviour. Geroski (2001) describes structural inertia as a consequence of 'rent displacement'; in other words, firms have already invested in existing activities that have a known amount of earnings. This suggests that structural inertia is a consequence of success. Hannan and Freeman (1984) claim that all firms are subjected to inertial forces, and that the selection process favours those firms whose structures are difficult to change due to this structural inertia. They argue that firms that change less are seen as more reliable and have higher accountability. In other words, those firms are more 'trustworthy'. The strength of inertial forces varies with age, size and the complexity of firms. Structural inertia is mostly internal to the firm; such as sunk costs in location, plant, equipment, and personnel. However, structural inertia is also found in the investments made in network relationships. Development of trust and routines takes time, so structural inertia increases with the age of a firm. Formalizing goals and standardizing patterns of activities, at a fixed location, stabilizes the organizational structure. This institutionalisation and standardization offer the advantage of reproducibility, but does increase resistance to change (Kelly and Amburgey, 1991). The reliability and accountability of a firm, in turn, increases as the firm gains experience and becomes better at the tasks it performs. According to the findings of Dobrev et al. (2003), path-dependent learning induces reproducibility of structure, which in turn makes the set of organizational actions of a firm more reliable. To put it differently, past experiences simultaneously creates survival advantages in the selection processes, and liability to inertia that endangers this survival. Which of the two assets determines the outcome of change depends on how the organization-environment dynamics unfold (Dobrev et al., 2003). Or, as Sorenson (2003) says, firms get better at making things as they gain experience producing them – they learn by doing. Through repeating processes, firms can continually improve their performance although, over time, these routines may become inert behaviour. The creation of knowledge within firms is an uncertain activity. Firms can handle such situations by developing internal procedures (habits) and routines when searching for possible solutions. A firm's habits and routines are mostly based on its successful behaviour in the past, and these habits are used as long as they are acceptably effective (Nelson and Winter, 1982). However, it is difficult and costly to unlearn once successful habits, even when they hinder future changes that would improve performance. Former success stories of firms often result in situations of structural inertia (Arthur, 1989).

Summarizing, large firms have strong structural inertia: larger firms have a lower speed of response to environmental changes (Hannan and Freeman, 1984). So, as a firm grows older and larger, structural inertia increases, making the firm slower in its responses and ability to change. At the same time, the firm becomes more

trustworthy and reliable; implying a greater likelihood of survival. Aging has advantages, such as improved capacities and more secure structural positions which tend to protect older firms from damage caused by changes in the environment. Hence, the life chances of firms improve with ageing (Carroll and Hannan, 2000). Or, as Ranger-Moore (1997) put it, inertia, while tending to hinder change, is itself the product of past change, commitment to sunk costs, entanglement in external relationships and so on; all of which represent organizational changes that increase inertia. Furthermore, change that modifies the visible mission of the organization (e.g. location or premises) undermines its legitimacy, based in part on reliability of performance (Delacroix and Swaminathan, 1991). Dobrev et al. (2003), see the concept of inertia as valuable for path-dependent learning. Past experience creates reproducible structures, which give firms reliability and survival advantages. The concept of path dependence is discussed more elaborately in the next section. On the other hand getting old has its drawbacks; lower response rates, higher costs and the possibility of becoming obsolete.

Thus, in population ecology, the environment acts as the source of selection mechanisms: the environment selects certain types of organizations to survive and others to perish based on the fit between their structural characteristics and the characteristics of their environment. Firms either fit or they fail. Although it is generally accepted that inertia inhibits firms from changing (Hannan and Freeman, 1984), the fact that, individual firms adapt to their environment is also accepted (Hannan and Carroll, 1995). In population ecology, the focus is on selection in populations; the principal, however, is not individual firms changing (Oinas, 1995). The key to connecting this idea to geography is understanding the range of processes that affect population dynamics, and not just the observed birth and death rates, since the world can be divided into spatially-delimited firm environments

4.3 Path dependency in explaining the survival of old firms

The concept of path dependency in its most familiar form is portrayed by examples such as the QWERTY keyboard and Betamax versus VHS video recording systems (Liebowitz and Margolis, 1995; Arthur, 1989). Other examples include driving in the left in Britain, and the rail gange in Russia and Spain versus the rest of Europe. Or, more on the firm level, the regional concentration of software producers in Silicon Valley and of pig and poultry producers in the northwest of Germany and the southeast of the Netherlands (Balmann et al., 1996). All of these cases share some initial advantage with increasing returns to scale and on network externalities. All of which result in so-called ‘locked-in’ situations. In other words, an equilibrium that may be inefficient, but which can only be abandoned at extremely high costs according to the founding fathers of path dependency: David (1985) and Arthur (1989). Antonelli (1997) argues that the

notion of path dependence can help in explaining the evolution of industries, and of individual firms. Firms are characterized by high levels of heterogeneity; including differences in size, age, organization, levels and methods of innovation, costs, market size, performance and, of course, different locations. This diversity in firm characteristics is persistent and in Antonelli's view, self-reinforcing due to path dependence. Antonelli considers industry groups and/or individual firms within a region, and not as Krugman (1991b), a core-periphery model concept of locational path dependency.
In this section, the concept of path dependence in evolutionary studies is defined and applied to the explanation for the existence of very old firms, with special attention to locked-in situations. Do old firms have a - spatial - path dependence and, if so, how did this path dependency either help or resist their long-term survival? Peneder (2001) argues that the models of Nelson and Winter, on evolutionary change, also provide a dynamic explanation of variety, arguing the continued reinforcement of diversity by means of the path dependent nature of behavioural routines, emphasizing the initial uniqueness of firms.

4.3.1 Path dependency; an introduction

In claiming path dependency, the most important argument used is that past events influence future events. This can be seen in two ways. Either those events that happened at an earlier point in time will affect the potential outcomes of a succession of events occurring later in time. Or, as the impacts of decisions made in the past persist into the present, they define the number of alternative choices in the future. In organizational evolution, path dependence is mostly seen as past directions enclosing or restraining directions for coming changes (Mahoney, 2000). According to Arthur (1989), path dependency has the following economic implications: future development can be predicted either, to only a small extent at specific times or, at other times, to a very large extent when the firm is locked in. In path dependency, a small event has long-lasting consequences (Antonelli, 1997). Once a path is taken, it can only be left at a high cost. In other words, path dependency is an allocation, under increasing returns, when actors choose between various options.
In a path dependent series of economic changes, the outcome is determined by historical, small events, or temporally remote events, including happenings that take place by chance rather than by systematic forces. Path dependence is thus stochastic. The dynamic process itself takes on an essentially historical character (David, 1985; Arthur, 1989). Putting it differently, firms cannot shake off the effects of past events, and do not have an unlimited choice of future options (David, 1993). Path dependence can be seen as a macroeconomic process, in which the actors have bounded rationality (limited information and time) on their actions and the timing of these actions. Furthermore, in path dependence, time matters,

because of the irreversibility of changes, which causes inert behaviour and affects the selection processes. Three forces create path dependence: irreversibility, indivisibility and the structural actions of actors. These forces provide a construct in which the effects of history on the behaviour of firms can be explained, in either their evolution or in their current status (Antonelli, 1997). According to Peneder (2001), with the survival of initial variety, firms in this way converge to unique (and optimal) behaviour.

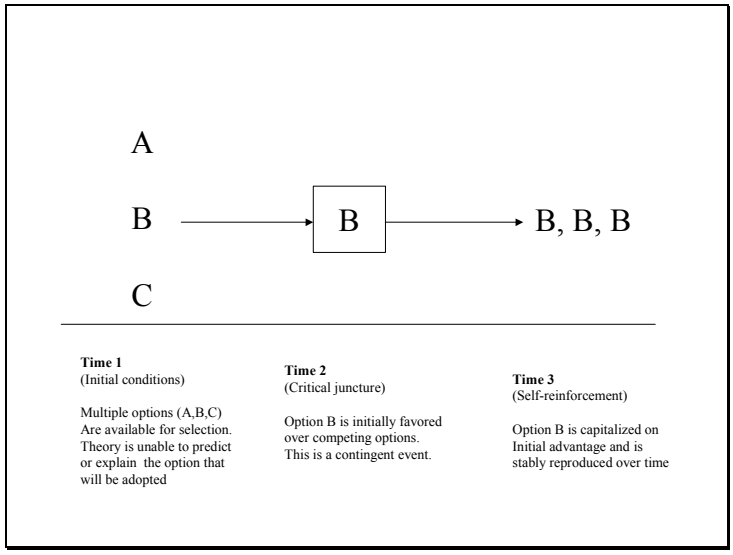
Although David (1985) and Arthur (1989) emphasize the role of increasing returns to scale and the role of network externalities in explaining path dependence, Balmann et al. (1996) show that path dependence can occur even when increasing returns to scale and network externalities are absent. Balmann et al. claim that the time structure of necessary reinvestments and sunk costs cause locked-in situations, which is the first step in the spatial application of path dependency. The example that Balmann et al. (1996, pp.163), give concerning: sunk costs as being of major importance when locked-in situations occur. The example is as follows: "Consider a natural monopoly held by firm F. This firm has a plant which cannot be sold, and therefore, has no better option than to stay in the market despite the attempts of a potential entrant E to take over the monopoly. F will hold the monopoly as long as sunk cost are significant. If all costs were variable, E could threaten to enter at any time." This shows that locked-in situations always have an inherently historical character. This historical character, as described in the example, can also be explained the following way: once the investment is made, it cannot be abandoned for a significant period of time unless the firm is willing to suffer extremely high costs.

Liebowitz and Margolis (1995) define three types of path dependence. According to them, three possible outcomes can occur in a dynamic process, depending on the initial conditions. In first-degree path dependence, this does no harm. The initial actions put the firm on a path that cannot be left without some costs, but it turns out to be best possible path. There is no direct inefficiency as a result. Second-degree path dependence occurs when the actors have bounded rationality, and the decisions made are not always the most efficient ones. Here the inadequacy of the chosen path is unknowable when the decision is taken, and at a later time one may discover that another path would have been more efficient. In other words, the sensitivity to initial conditions led to undesirable outcomes that are costly to change. However, the existence of a better, alternative path was not known at the time of the initial decision. Third-degree path dependence is also sensitive to initial conditions, leading to undesirable outcomes, but now the outcome could have been known at the time of the initial decision. Nevertheless, the path can be changed with costs. The first type of path dependence is an example of past-dependence but with no direct undesirable outcomes. The second type also has the same dependence on the past; however, due to bounded rationality, the most efficient path was not chosen. The third type is also dependent on initial conditions but the error could have been anticipated and avoided.

To summarise, history matters, or the past influences the future. Time matters because the outcome of any change in market dynamics is dependent upon the character of the initial setting of the market and the behaviour of actors at any point of time (Antonelli, 1997). According to Mahoney (2000), path dependency is more than simply tracing a given result back to a specific set of historical events but also involves showing that these specific events were themselves contingent episodes that cannot be justified solely on the basis of preceding historical conditions. Consequentely Mahoney (2000) claims that there are two types of path dependent sequences. The first is the 'self-reinforcing sequence', characterized by the formation and long-term reproduction of a given institutional pattern. 'Self-reinforcing sequences' often demonstrate what economists call 'increasing returns'. With increasing returns, a once adopted pattern, delivers increasing benefits by repeatedly adopting the same pattern. Thus, over time, it becomes increasingly difficult to transform the chosen pattern or select other, previously available, options, even if these options would have been more efficient in the long run. This is comparable to path dependency as described by David (1985) and Arthur (1989). The second type are the so-called 'reactive sequences', which are chains of temporally ordered and causally connected events. This kind of sequence is reactive in the sense that each occurrence within the sequence is a reaction to an earlier event. Thus, each step in the chain is dependent on prior steps, similar to the sunk cost approach by Balmann et al. (1996). With this kind of sequence, the final event in the sequence is typically the outcome under investigation, and the overall chain of events can be seen as a path leading up to this outcome. In this specific research, the survival of old firms.

In Figure 4.2, an example of a reactive sequence as given by Mahoney (2000) is shown. On the basis of initial conditions, present at a certain time, the reason for adopting option B cannot be predicted or explained. In a reactive sequence, given the initial conditions, one could theoretically 're-run' history many times, and there is no reason for believing that option B was the optimal choice, or any reason why any of the other options (A or C) should be adopted over option B. The initial adoption of option B is a contingent event. As Figure 4.2 implies, once B is contingently selected, it is stably reproduced over time in the future. The options that Mahoney (2000) give are either first degree or second-degree path dependencies in the definition given by Liebowitz and Margolis (1995). This is also the most likely path dependency for old firms in this specific investigation. Since history matters, in order to cope with the uncertainties and incremental character of learning processes, firms develop routines that are very durable over long-term periods and this, in turn, will tend to establish path-dependent 'learning trajectories' (Maskell and Malmberg, 1999). Over time, generations of embedded knowledge and other sunk costs will solidify the once chosen investments, and thus limit the choice of options the firm can take in future, as well as explain the choices made in the past (Dosi, 1990).

Figure 4.2: Path dependency when initial conditions favoured B



Source: Mahoney (2000)

4.3.2 Path dependence in organizational ecology and industrial organizations theories

The condition of a particular industry at any time is dependent on the condition of the industry in the period previous to the current period. The description of the transition between these periods is, according to Nelson and Winter (1982) the application of evolutionary theory. In this way path dependence can be incorporated in the models of organizational ecology theory and industrial organization theory. Antonelli (1997) described this as the dynamic processes in which the transition from one state to another depends on earlier conditions. The complexity of outcomes from the interactions of actors (firms), that are fully embedded and localized in the structural characters of the system, are still influenced by evolution through to a variety of structural actions (Antonelli, 1997). This is a first-degree, and reactive sequence, path dependency as explained in Section 4.3.1. But is this accurate? As Forday (1997) puts it: Do real-world competitions between different technologies show a lock-in to an inferior technology because of initial and trivial circumstances? Is path-dependent selection likely to occur, rather than natural selection, in the real world? With natural selection, the fittest survives (see Section 4.2.5). By contrast, in path-dependent selection (except for the third-degree, see Liebowitz and Margolis, 1995), the observation is that the sequence of choices determines what is the best fit instead of an objective best choice (Forday, 1997). However, these processes do not have to be mutually exclusive. Mueller (1997) finds that most industrial groups experience the following scenario, which is similar to the density dependence model. In a

Schumpeterian way, the industry starts after a significant innovation. Then, imitator-firms, often in great numbers, follow these successful innovators. During this start-up phase, output expands fast and prices fall. This implies product improvements. At some point in time, the market starts selecting the best fit, that is the best model designs and firms start to develop the best ways to produce this product. The firms that have chosen the best 'path', either in product or process, survive and the others exit. Following this so-called 'shake-out period', the dynamics become increasingly stable and the numbers of entries and exits no longer change dramatically. This mature phase of the product life cycle may go on indefinitely, so for the survivors that have chosen the right path there is no reason to exit.

Antonelli (1997) developed the following argument. The parameters of the market are not exogenous or static, but are the outcome of the past behaviour of actors as well as the dynamic properties (density, legitimation and competition) of the market itself. The notion of path dependence can be seen as a general framework in which the evolution of industrial economics is explained. This explanation goes beyond the standard evolutionary approach. Since, in reality all firms are different in terms of size, age, costs and other capabilities, a set of structural factors in a given environment can be assumed to act as selective tools. This explanation, based on the idea of path dependence, provides a more general framework which takes into account both the decisions of the actors and the structural factors in an interactive way. The past behaviour, choices and decisions of the actors can account for the possible adaptation, learning and creation of inert routines by the firms in response to the changing structural factors of the environment, now and in the past. Or, as Antonelli puts it: "Hence, path dependence provides a framework to understand and model the effects of historic time on the behaviour of agents which however are able at each point to modify their evolution" (pp.669). In this way Antonelli argues that "the growth and diversification of economies of scale, economies of scope and learning processes play a central role in explaining the evolutionary path of firms through markets, products, countries and technologies so as to explain the variety of specialization profiles and the variety of processes of growth. The interaction between different learning processes, sunk costs and industrial structures play an essential role in assessing the evolution of the performances and the strategies of firms and hence the persistence of economic profits and innovative activity on the long run" (pp.669).

4.3.3 Age, inertia & historical lock-in: path-dependent?

Path dependence illustrates the mechanism that explains the way in which the legacy of an industry or an individual firm can influence present choices and outcomes (Scott, 2001). In other words, small events that happened long ago still have an influence on how a firm functions in the present. Within a firm, several generations of firm-owners or managers have made decisions that determined

choices (e.g. about investment in plants, or technology), and will determine other choices in the future (Scott, 2001). David (1993) describes this as a path between three points, the present we live in, the point of departure and the series of occurrences connecting the point of departure, with the present. If we want to know the meaning of the present we need to know the history. Path dependence can be seen as a form of inertia. Path dependence implies a level of irreversibility, persistence and durability either in a technical, institutional or economic sense. This path dependence inertia is caused by two factors: first, the level of investments in the old technology, materials and location (sunk costs); second, the uncertainty that firms have about the behaviour of others (Forday, 1997). Mueller (1997) also acknowledges inertia in firms, but from the buyers and consumers side. Each time a product is bought that has a greater utility than costs, there is a high chance that this will be repeated. As long as the purchase of the product maintains this advantage, the consumers and buyers will repeat this action and it will become a habit, i.e. one buys the product almost without thinking. In behavioural psychology this is called the result of operant conditioning (Mueller, 1997) or, as earlier described, this raises the level of legitimation, that can be explained as follows.

Antonelli (1997), as described earlier, sees path dependency as having three factors: irreversibility, indivisibility and the structural actions of actors. Irreversibility can be defined as the difficulty of changing a given behaviour or choice - the costs that have to be incurred in any attempt to change a commitment to a given behaviour or decision taken at an earlier time. All sunk costs are irreversible. Indivisibility leads to economies of scale, externalities and economies of scope, i.e. transaction costs. For firms with indivisibility and irreversibility the outcome is the firm's reputation (Antonelli 1997; Peneder, 2001). Mahoney (2000) labels this the particular circumstances that caused a tradition to be started, which is reproduced by social patterns of habits. This can all be recognised as first-degree path dependence. According to Liebowitz and Margolis (1995), all economies or firms have long-lasting characteristics that reflect first-degree path dependency. What is today, is in part a consequence of what was and has been done yesterday. This is called the inheritance of capital stock, language, customs, laws, grudges and skills. Heffernan (2003) refers to these habits as response patterns that evolve out of behaviour and methods used within individual firms as well as out of the behaviour of customers. Liebowitz and Margolis (1995) give the following example. A firm with fixed assets will continue to use an inferior technology when the average costs of the old technology are lower than the costs of a new technology. Under such circumstances, the old firms might be considered 'locked-in' to this inferior, but still more profitable, technology, exhibiting a first-degree form of path dependence. In the longer term, however, this old technology might become part of their culture, reputation and corporate identity: the lock-in is here to stay. In this way the endowment of one generation is the bequest to another

generation (David, 1993). Mahoney (2000) argues the same, however calling it legitimation. Legitimation explains that, once a given tradition is chosen, the repeating of this tradition will increase the legitimation, even if other possible customs would have been more acceptable in the longer term. Increasing legitimation, as described in Section 4.1, is important in increasing the likelihood of survival. This results in a cycle of self-reinforcement. The tradition that is chosen at the start creates a standard for legitimacy, and is for this reason reproduced. This by itself creates even stronger forces for reproduction of the tradition (Mahoney, 2000). From this perspective, the persistence of traditions that seem less functional and less supportive of elite interests can also be explained. Or, as Peneder (2001, pp.40) describes this, "the initial variety of firms is preserved along its particular path of development, which is shaped and constrained by the firms' specific set of critical and irreversible choices made in the past." The longer a firm exists, the stronger this kind of inertial forces become which can be traced back to earlier events through path dependence. Old firms get locked-in by their reputation, the legitimation of their product, and identity. Over generations this becomes a self-reinforcing process within the firm.

4.3.4 Spatial path dependence

According to Forday (1997), the initially chosen path is likely to be continued, rather than changing paths. This is due to increasing returns, especially through learning by doing, learning by using or some other similar dynamics. In this sense, it can be said that localized learning can be seen as an irreversible, path dependent mechanism (Malmberg and Maskell, 2002). If this spatial process can be seen as path dependence, how about other spatial processes, for example the choice of a firm's location? Is this path dependent?

Liebowitz and Margolis (1995, pp.217) use the following example of the choice between alternative locations where there is a spatial agglomeration economy under circumstances of increasing returns. They assume that payoffs depend on the number of prior adopters, and that adopters have some knowledge of the increasing returns of payoffs. "…firms can locate at one of two geographic sites, A or B. As a result of agglomeration economies, payoffs increase with the number of firms that choose each location. These are of course, the increasing returns conditions that are held to bring about lock in. But there are two questions that arise here. First, given the pattern of returns, how does either location attract any firms, given that there appears always to be a return to delay? And second, which location prevails? The solution to the problem of getting any firm to locate at any location comes fairly readily. It is landownership. In either a Ricardian or Von Thunen framework, any unique advantages of a particular location accrue to the landowner as site rents. Under conditions that are reasonably likely, owners of land can expect to extract the rent created by any locational advantages that eventually attach to the land. Where agglomeration economies are fairly localized, a pioneer firm can appropriate the rents that result from agglomeration economies by acquiring all the land in the close

proximity to the initial plant. Alternatively, ownership of other unique resources, such as transportation facilities, may constitute opportunities to internalise the agglomeration economies. Someone must capture and redistribute these rents, or else payoffs like those in the table provide no way of getting started: everyone want to go last. … Landowners have incentives to offer adequate inducements to potential early entrants. They can offer long-term leases at lower rents, or subsidize infrastructure, or even pay direct subsidies to early arrivals. Rewards offered to initial entrants will be recaptures out of the larger surpluses generated by subsequent entrants. The exogenously determined rewards … thus can be understood as incorporating artificial rigidity of incentives or inflexible prices of productive inputs and outputs." This view, however, considers the firms to be neo-classical actors, and ignores all notions of individual behaviour and bounded rationality as described by Simon (1955) and Pred (1967).

Considering the earlier described three degrees of path dependency, this can be read that, for most firms in this particular research, location is a form of either first-degree or second-degree path dependence. Most firms have the tendency to locate in the vicinity of the founder's home, since this is the region they know best. New entrepreneurs look around for a location in their familiar, often immediate local environment and make a fast decision, rather than carry out an extensive search for an 'optimal location' (Simon, 1955; Stam, 2003). Hence, the initial location decision is a non-economic one. For example, the Dutch Ministry of Economic Affairs (1998) found that more than 60% of current firm start-ups were begun from the entrepreneur's own house. In the founding era of old firms, before 1850, this can be expected to have been even more so, due to the prevalence of crafts production, and slow and difficult transportation (see Chapter 3). So, firms choose their location 'by chance' but this could still be to be the optimal choice through first-degree path dependence. Or second degree, if the firm's location was first chosen more by bounded rationality and, in the course of time, they discovered that another location might have been more efficient and that relocation (different path) was an option, however, only at considerable costs. In the same way, firm relocation can also be seen as second-degree path dependence, despite substantial costs having to be incurred to move to another location (or path). Nevertheless, in most cases, the firms could not have anticipated the need for relocation such as due to growth of the firm, at the moment of the initial location decision. Furthermore, some firms see the need to relocate, realising that other locations would be more beneficial but due to the high cost of changing paths, they are 'locked-in' to their initial location. This is also related to legitimation or identity as described in Section 4.3.3. If a choice acquires a 'momentum', in other words the opportunities for change are more dramatic than the opportunities for remaining the same, this is called a negative lock-in. According to Håkansson and Waluszewski (2002), 'negative lock-ins' are caused by the interrelatedness of technology, the compatibility of systems, the economies of scale and the irreversibility of investments. However, not all lock-in situations are without question negative. Peneder (2001) finds lock-in effects that are critical in the preservation of initial

variety, which is important for firms to get positively selected in organizational ecology. Through ‘positive lock-ins’, individual firms generate functional differentiation in an evolving system. Identity can be created by positive lock-ins. In contrast to what Krugman (1991b) argues, not all firms converge towards the same type of behaviour. Krugman assumes that spatial path dependence is caused by transportation costs and economies of scale due to agglomeration economies (comparable to the example of Liebowitz and Margolis, 1995). This path dependence creates core and peripheral regions with a geographical concentration of specific sectors. However, in this particular research a specific set of individual firms are investigated, and regardless of whether their behaviours correspond they are all treated as individual cases with their own specific behaviours. Path dependence on the level of the individual firm can occur through, for example, technological consideration as in the examples of Arthur and David, but also as a result of the choices or adoption of a set of rules (Balmann et al., 1996; Heffernan, 2003). These rules can become habits, tradition and identity, and create ‘positive lock-in’, including the location, vicinity and the region in which the firm is located.

4.4 Spatial implications of the evolutionary view on old firms

Theories concerning Regional Science can take two different approaches towards the firm and the environment. First, one can infer the behaviour and functions of firms from the environment in which the firms are located. The spatial environment is, in this view, of overriding importance for the formation, behaviour and performance of the firms. Here, little space is left for the firm to influence the environment in which it is located. Second, one can see the environment in which the firms are located as the result of the firms' behaviours and activities. In this view, the firms have more freedom of choice. There is close attention to the firms’ decisions and structures and the influence these have on the spatial environment. In the current research the first approach is chosen. It is assumed that the survival, the 'aging' of firms, is to a large extent specified by their environment. The spatial environment in which the firm is located, and how the old firm behaves in this given environment is presumed to contribute to the chances of survival. Variables that can determine this survival include: the density in the location region, the size of the firm and the size of the firm's market, transportation and firm accessibility, and the size of the competition.

4.4.1 Variables in long-term firm survival

In population ecology, the firm population is studied on the level of entry and exit of organisations in specific industries. This study, as such, does not study old firms in the Netherlands on such a population level; however, the models and propositions resulting from research in the field of population ecology do provide

interesting approaches to investigate old firms, and also several industrial sectors at the same time. With the long time frame of this investigation, it is not possible to acquire entry and exit numbers for the population at stake (see Chapter 2 for a more detailed discussion). Nevertheless, it is interesting to investigate just the survivors, and see whether they have the characteristics that one might expect from theory. Which firms have a higher survival probability, why are there these differences in firm survival, and what are the population consequences? As argued earlier, the interaction of the firm with its regional environment is a crucial determinant of firm behaviour. This regional environment can be described as linkages of customers, suppliers and deliverers, local and regional competition, and more generally a network of relationships with usually a strong spatial dimension. New explanations of these phenomena have focused on the concepts of history dependence, chance and learning regions, which are familiar terms in the field of evolutionary economics (Van Wissen, 2002). Now that the theories and instruments have been discussed in some detail it is important to develop the theoretical concepts into useable construct within a spatial context. After all, this research wants to find spatial factors that influence the long-term survival of firms in the Netherlands. Therefore, the goal is to find the critical variables affecting survival, with the necessity to derive measures for these variables in a spatial context, in other words the location of the firms.

Spatial adaptability and inertia

Caves (1998) and Ranger-Moore (1997) argued that the survival of mature firms is determined by the degree of depreciation or obsolescence in their original cumulative learning abilities. This means that if the firm's abilities for learning become out-moded or old-fashioned it will likely perish. When these learning abilities reduce over time, this will also decrease the likelihood of survival of the firm. But what are these learning capabilities? The learning abilities of firms can be seen as the way firms adapt to changes in their environment. The environment of the firm, according to Walker (1975), includes all factors not in control of the firm. These include the actions of suppliers, buyers, government departments, and the whole range of events that comprise the general economic and social conditions in the society in which the firm operates. Vaessen (1993) uses Walker's view in his research 'Small business growth in contrasting environments', although Vaessen clearly makes a distinction between the 'business' environment and the 'spatial' environment. Vaessen (1993) adds the location of suppliers, customers and other actors relative to the location of the firm. This is the approach that will be applied in this study. The learning process that copes with the spatial dimension is referred to as the spatial adaptation (Stam, 2003). The term spatial adaptation is used for those types of adjustment in which a deliberate effort is made to cope with the spatial dimension: for example, the price of inputs from various sources, the sizes of demand in different areas, or the cost of transportation changes. As a result of

environmental changes, a firm may adjust its pattern of good linkages to improve its profit record. One form of spatial adaptation could be locational adaptation; in this case, management decisions directly affect the location of the firm's production, for example through the building of new plants or the expansion or closing of existing ones (Walker, 1975), or by making adjustments at the current site. Watts (1975) agrees with this view: the survival of a firm in the long run depends on the ability of the firm to extend or intensify its market area. Vaessen (1993) agrees that a firm can learn to survive in contrasting environments through the extension of its market to a national, or international, level. In the case of the survival of a firm, the legitimation can be seen as a certain amount of embeddedness in a bounded environment or region (Stam, 2003). The more institutional links a firm has in its own region, with for example suppliers, customers and governmental departments, the more it is generally accepted that the firm 'belongs' there, and the likelihood for survival will increase. With this increased embeddedness of the firm in its location over time, the firm develops inertia to change, and locational adaptation in the form of relocation becomes less likely since this kind of change would be very disruptive and reset the liability-of-newness-clock (Hannan and Freeman, 1989).

From the above, it can be inferred that long-term surviving firms show higher - locational - inertia since these firms have not suffered from loss of competence in the longer term (Delacroix and Swaminathan, 1991). In other words, inert behaviour can be seen as a measure of long-term success. This argument, however, does not fit with the population ecology tradition but rather fits the organizational learning point of view, the position adopted by March (1991). March argues that organizational change may be beneficial in the short run but harmful in the long run. The ability of firms to change is limited by the structural conditions, both internal and external, in which the firm is embedded. This argument is opposite to the argument that firms are adaptive to environmental shifts and are supposed to be able to adapt, as illustrated in a long-term study on the evolution of firms in the US automobile industry (Dobrev et al., 2003). Kelly and Amburgey (1991) argue that structural inertia varies with a firm’s age and size. Since old firms have had time to formalize relationships and standardize routines, the result is increased structural stability. This increasing stability, in turn, increases resistance to change, so inertia also increases with age. Consequently, the probability of change in firms declines with age because locked-in learning and inertia result from past changes and experiences (Dobrev et al., 2003). In transferring this observation into the research of old firms, it can be expected that old firms may have demonstrated adaptative behaviour in the early periods of their lives but with the increase of age, these firms display increasingly inert behaviour, which in the spatial perspective comes back to less locational adaptation.

A relocation decision, however, is not a common occurence in most firms, since there is much uncertainty attached to the relevant procedures (Townroe, 1979). The

first of these is the cost of finding and developing a new site. Frequently, large expenditures are required to collect the necessary information from a thorough search, as well as in the purchasing and development of the land, and the building of the premises or factory. Costs will be incured as well from the interruption to production (especially in the case of relocation) and the time spent in bringing a new site into commercial production. These costs tend to discourage firms from considering new locations, especially relocations, unless the firm is facing major problems at its existing site. Thus sunk costs also encourage locational inertia. A second aspect is that locational decisions inevitably involve a large element of risk or uncertainty. Such decisions require the provision of fixed capital with several years of payback while the future conditions facing the firm cannot be know for certain (Pred, 1967; Walker, 1975; Townroe, 1979). With locational adaptation, the liability-of-newness clock might also be reset (Amburgey et al., 1993). The concept of adaptation can be rephrased as locational change reflecting both changing environments and the reaction of firms to these changing environments.

The focus on individual firms is important in identifying some of the key aspects of locational adaptation. For the purpose of understanding the dynamic development on an industrial space economy, however, it is crucial to transfer these insights to the aggregated situation (Walker, 1975). As Wintjes (2001) finds in his research on the embeddedness of foreign firms in the Netherlands, the involvement and commitment of firms to their local environment is based on each individual firm's actions as an economic actor based on its specific background: culture, tradition and identity. If the firm shares 'local habits, culture and knowledge' with its environment, the region and the firm become more committed to each other. Also Vaessen (1993) argues this, that the firm influences its own local environment, by improving the circumstances in the environment. An example is infrastructure since it helps the firm to function but, on the other hand, also increases the embeddedness of the firm with its local environment. According to Wintjes' findings, the activities of the firms and the characteristics of the region do influence each other mutually over the course of time. The firm cannot be seen without considering the development of its regional economy. Also Malmberg (1997) find evidence for this, in Sweden. According to him, most regions retain their once-adopted specialization, indicating that an embeddeness that grew over time can sometimes last longer than the resources on which this specialization was once built.

From the above, it can be anticipated that old firms will be generally large, and exhibit inert behaviour. This inert behaviour can be internal to the firm (product, culture, routines etc.) but also in locational adaptation. Old firms are not expected to change locations after a long period of 'embeddedness' at their current location, having become intertwined with their regional economy.

Density and competition

The density of firms in a region is also important for the chances of survival. Firms founded in periods of high density are expected to have higher age-specific mortality rates. Resources can be scarce in high-density periods, conditions that could cause mortality in less strong organizations. In this framework, the density variable is an incomplete way to measure evolution since not only the number of firms is useful in measuring evolution: size and location should be included (Swaminathan, 1996; Carroll, 1997). The higher the density is in a certain area, the more competition there will be in this region, but also there will be increased levels of legitimation. Larger firms will have more power to reach the resources and, because of economies-of-scale, a better competitive position against the smaller firms (Aldrich and Pfeffer, 1976). This is also applicable for the location variable. When a firm is situated in a better location (for example, with greater accessibility and closer to resources and customers than its competitors in a high-density region) it clearly has a higher likelihood of survival.

In most cases, the birth of a firm takes place in an urban environment, whereas the growth of firms often takes place in a suburban environment. It is found that economic dynamism, reflected in high birth and death rates, is generally higher in urban regions than in rural ones (Pellenbarg and Van Steen, 2003). Or, as Oinas (1995) put it, core regions 'select' contact-intensive industries with high requirements in terms of the quality of labour. Peripheral regions on the other hand, due to their specific environmental conditions, tend to 'attract' firms with production or service functions requiring a less-skilled low-wage workforce. In this last explanation, the nature of the industry is important. This will be elaborated further in the subsections below. Van der Knaap and Van Geenhuizen (1988) also argue that the role of the environment is important for firms, especially for the level of innovativeness. Firms located in a metropolitan or urban environment have advantages due to 'localized learning' (Malmberg, 1997). Firms with higher levels of innovations are better able to 'fight' competition, and hence have a higher likelihood to being positively selected. The competitive environment is seen as being more local than institutional, in other words the distance to the competitors is of importance (Bigalow et al., 1997). For survival, the competitive environment of firms has to be investigated in terms of close proximity. Increasing the density in the region of more or less the same kind of firms increases competition. Furthermore, a firm itself becomes more competitive (and positively selected) with increasing firm age and size (Rodríguez et al., 2003). On the other hand, the stiffer the competition in the firm's environment, the greater the need for improvement (Walker, 1975). However, due to structural inertia, it is not always possible or desirable to change internal or external habits and routines.

From the arguments, old firms can be expected to have survived strong competition, in order to obtain legitimation for their product. In the long term, this can be traced back to earlier times when the firms would have had sufficient

competition at a regional level. Furthermore, old firm are expected to have flourished more in urban areas through gaining positive influences of localized learning, economies of scale and greater innovativeness.

Size of firm and market

Empirical evidence indicates that firm mortality tends to decline with increased size, and therefore large firms are expected to be less vulnerable to the risk of failure (Baum and Oliver, 1991; Freeman, 1981). It should be recognized that there is some evidence that the concentration of resources in fewer large organizations is increasing, and therefore that these larger firms will dominate many aspects of current economic life (Aldrich and Pfeffer, 1976; De Jonge, 1969). As stated earlier, larger firms have more power and better economies of scale, and therefore will have greater survival chances. Another way of reasoning is that large firms started as small firms, and have managed to survive until now. Due to their long-term success, it can be anticipated that older firms are, in general, larger than average (Dunne and Hughes, 1994; Rodríguez et al., 2003). Size, in this research is measured by number of employees, although there are many other ways of measuring firm size and growth (Brouwer and Henrich, 2001). It is chosen to measure firm size in number of employees, since this information is freely available. The growth of firms, also measured in terms of numbers of employees, can represent value creation for the firm or can indicate organizational strengthening (Rodríguez et al., 2003). They argue that the growth of a firm contributes to its survival and competitiveness, however, when a firm age, its growth diminishes. Dunne and Hughes (1994) tend to agree, although they refer to this as that the survival of firms is mostly related to past growth. Firms display slower growth over time, since as firms grow older they generally tend to be larger, develop formalization, routines and internal consistency (inertia), which reduce growth rates (Entrialgo et al., 2001). Growth, or this proven survival or fitness, could be due to location, product differentiation, lack of competition or innovation, etc. Increase in market size is also an important variable. An increase in market size, such as from a regional market to a national or international one, can be a positive determinant in long-term survival (Vaessen, 1993). From a spatial perspective, this increase in market can be achieved by delocalisation of production and distribution or by opening branches in other parts of the country or abroad (Mariotti, 2004), or by structural adjustment at the initial location (Knaap and Van Geenhuizen, 1988).

Gibrat's stochastic model from 1931 is often used to explain the growth dynamics of firms (as quoted in Rodríguez et al., 2003). According to Gibrat's law of proportional effects, size does not influence growth. Growth is described as a phenomenon resulting from several factors, but not directly linked to the size of the firm. Jovanovic (1982) finds that small and young firms grow fast although, with the evolution of the industry, growth shows an inverse relation with firm age and

size. This is due to the 'learning process' firms go through with aging. According to Jovanovic, larger firms are more efficient than smaller ones and this is the result of the selection process, in which efficient firms grow and survive, while inefficient firms stagnate or exit the industry. Since this process takes time, larger firms are inherently older, which according to Jovanovic implies that there is a positive age-efficiency relationship.
Older firms can be expected to be, on average larger (in number of employees) and have a larger product market. Furthermore, while it is expected that old firms have experienced periods of growth in numbers of employees, this does not need to have been in their recent past.

Founding circumstances

The mortality rates of firms in a particular cohort may depend not only on organizational age, but also on the interaction between the firm's age and initial conditions such as environmental conditions at founding. In adverse environments, weaker firms are eliminated quickly and the survivors are hardened by battle. Firms founded in adverse environments thus experience higher initial mortality rates. However, beyond a certain age, the survivors among such firms have mortality rates that are lower than firms founded in less adverse environments. Perhaps the adverse founding conditions eliminate the weaker firms more quickly, leaving behind a surviving cohort of stronger firms which results in a population with lower mortality rates (Swaminathan, 1996). This reflects Walker's (1975) conclusion that the fittest firm in the struggle for life will be the survivor. How can these adverse environments be evaluated? What are the elements of an environment that make it adverse in comparison with other regions? Amburgey and Rao (1996) say that ecological research on mortality overlooks how firms are embedded in social networks. This assumes that social networks are also regional networks, because the contacts a firm has in an industrial districts are mostly over a short distance (Markusen, 1996). This could be one argument: if a firm has a less strong social network, it has less chance of survival. This is because it will take more input to get the same output. If one looks at the study by Vaessen (1993) 'Small business growth in contrasting environments', it is clear that it is very possible for firms to start up in environments that are less than ideal and succeed in these environments. Further, these firms, when they survive the first years, have enhanced chances of survival in the long run. This is because these firms have had to cope with longer distances to customers and suppliers from the beginning. Such firms start with a much broader market or a much larger network in geographical distance than firms that start in cities with a very regional market. Vaessen concludes that the success of firms is a function of both the way they behave and the economic environment in which they exist. Furthermore, Schot and Rip (2003) point out the importance of time in the founding conditions. They found that there are certainly periods in history in which it was harder for new entrepreneurs to start

up firms, as was discussed found in Chapter 3. However, in practice, research on firm specific-founding conditions can only be done in very general terms, due to the lack of firm-specific information caused by memory-gap. The founding conditions of old firms can be studied as follows; the era in history, economic conjuncture, and whether or not the firm started in an urban or rural environment. Old firms are expected to have emerged in a period of economically conjuncture mostly in rural environments.

Sector, networks, innovation and identity

Industry-specific survival has been researched intensively, including at the empirical level. For example, Van Kranenburg et al. (2002) did not find significant evidence for industry-specific survival in their research on daily newspapers in the Netherlands. Siotis (2003) carried out a study in Spain and, although with a shorter time frame, also did not find sector specific differences. However, Melarba and Orsenigo (1999) did find survival differences at the sectoral level, although they only tested different high-technology sectors. Esteban (2000) also found sectoral differences for survival within the European Union, but only when they were region specific, and the influence was of minor importance. Dejardin (2004) finds that sectoral and cross-sectoral dimensions are relevant in the interrelationship between firm birth and deaths. Industry and sector seem to be determinants in the evolution and survival of firms. In the sample of old firms studied here, mostly the manufacturing sector is represented. Or, at least, most of the firms started in manufacturing although sometimes, over the course of time, due to product-differentiation, they are now in another sector. So instead of making a product, the firm is now trading in that product.

Entrialgo et al. (2001) find that older firms are generally larger and, due to old age and larger size, these firms are associated with increased inertia which, in their findings, is negatively related to innovation in firms. However, Huergo and Juamandreu (2004) find evidence for the opposite. They look at the impact of a firm's age and on - process - innovations. Overall, the innovative activity of firms does not diminish with the passing of time. However, when also sector differences are incorporated, they find that within manufacturing the opposite is true and that older firms tend to show fewer innovative events than those in other sectors. These findings indicate that innovation has an effect on survival. Innovation activity is generally found to be related to the size of the firm. Schumpeter (1942) hypothesized that large firms have an advantage in innovations. Since new firms are, on average, smaller than established firms, they thus have a disadvantage in this respect. Hence, it can be assumed that older firms (that are in general larger) will be more innovative. However, Brouwer (1998) found that there is no direct empirical evidence of a relationship between firm size and innovativeness. She finds that when innovation is measured in direct innovation counts, small firms are much more innovative than large firms. However, the same could be argued for old

firms, since they have been innovating for a longer time and so, in direct innovation count, they would have higher innovativeness. According to Brouwer, innovativeness is sectoral-dependent. In highly innovative industries, innovation rates are higher in larger firms due to capital intensity, product differentiations, unionisation and concentration.

Networks reflect the patterns of geographical industrialisation as described by Taylor (1995), or embeddedness: the contingent nature of economic action with respect to cognition, culture, social structure and political institutions. This view regards social forces as a sort of 'cultural software' that determines the behaviour and decisions of actors (Granovetter, 1985; Zukin and DiMaggio, 1990). Networks are the sunk costs of a relationship, which provide a basic condition for trust, stability, reduction of opportunism, and reduction of cognitive distance. Conversely, sunk costs become a threshold for change (Dahlgren, 2002). These investigations indicate that networks are important in some way for survival. As Oinas (1995) put is, it is easier to gain legitimation if you are similar to the ones that you frequently deal with in a network relationship. Furthermore, localised learning is, as Malmberg (1997) describes it, the role of the location of the firm in its process of learning. He indicates that firms in the right type of 'local milieu' (agglomeration) benefit from the spatial proximity of other firms. He argues for an interaction between the existence of industrial agglomerations and the durability of patterns of regional specialization. This localized learning of many firms clustered together is, according to Malmberg, in a Hägerstrand way a diffusion of positive influence on the level of innovations and technological changes. Firm learning is complex, multifaceted and history dependent. The key people in the firm are all involved in their local or regional communities and, as a result, receive all kind of economic, sociological and political information that may stimulate or limit the firm's development and thus influence its strategy and survival and vice versa (Julien, 1995).

Older firms have existed long enough for the routinisation of internal and external linkages, and this has a direct influence on their culture, reputation and identity. Old firms have been so long in the same place, with more or less the same product, market or process, that changing these would hurt legitimation and with this its reputation and identity (Amburgey et al., 1993). This can perhaps be called the 'liability of changes'. Firms (or an industry) that have a particular product have, with the passing of time, created legitimation, accountability and reliability with this product, reflecting their culture, reputation and identity (Kelly and Amburgey, 1991) and so creating an embeddedness in their - spatial - environments. Any change that directly disrupts this embeddedness would reset the clock. Such firms remain the same through structural inertia, and reduce failure rates by continuity. Dobrev et al. (2003) bring into this argument the view that past experiences add to experience and result in structural reproducibility that develops inertia. Also Baron (2004), who researched organizational identity in organizational ecology theory,

points out the necessity of a 'authentic identity' for survival, and acknowledges the increasing inertia which is induced by this identity. According to Baron, the identity of a firm consists of "a set of interrelated rules, assumptions, beliefs and premises that lead to prescribed patterns of behaviour" (pp.3). Changing the firm's identity might be very disruptive for accountability and reliability and a firm may for this reason adopt inertial behaviour. "Changes in the identity are highly destabilising with regard to employee turnover, survival and market performance" (pp.6), The particular basiss of identity are product, legal forms, industry and geographical boundaries, as well as product market, corporate culture and structural distinction (Baron, 2004). These identity factors are very robust to changes in the organizational population. Furthermore, the identity of a firm is created early in its evolution and is in this way path-dependent. Once an identity-creating-factor is chosen, such as a product, premises or habit, it will be hard to change the path that was chosen. Furthermore, a firm will resist changes to these identity-factors, since they determine the firm's level of trustworthiness.

Van Geenhuizen and Nijkamp (1995) developed an evolutionary perspective on firm survival, the 'company life history approach' that has a specific attitude to strategic change. They are of the opinion that firms need to be strongly adaptive in order to survive. Normally, firms adjust to their environment in an incremental way, however, sometimes there is a need for more radical change, such as a merger, new technology, new product, or relocation. This form of company history analysis, in their particular research, is more macro - on the aggregate level, although attention is paid to the individual firm. Nevertheless, in our investigation of old firms, the group of 'oldies' is also studied in the light of the development of the Dutch economy. Van Geenhuizen (1998), furthermore, argues the importance of period effects in explaining firms' dynamics. However, evolutionary economics emphasizes that firms develop, stabilize and follow routines. These routines may change over time, but mostly they function as stable carriers for knowledge and experience. This causes inertia. Spatial lock-in can be the result of the spatial formation of newly emerging industries as an evolutionary process, in which there is a spatial concentration of increasing returns (agglomeration economies). According to Boschma and Lambooy (1999), the evolutionary school of thinking is divided between long-term evolution of the economic system in terms of dynamics (new variations), and stability (development without any change in the parametric configuration).

Overall, old firms can be expected to show stability in identity, however, they also adapt in some ways, such as market, size and product-differentiation by process innovation. While old firms will have an 'unchanged' identity, at least for the outside world, underneath minor adaptations in processes might well have taken place over time.

4.4.2 Hypotheses for testing

Since the age and size of firms have an influence on their survival chances, it is to be expected that the surviving firms are generally larger than the 'average' firm. This also raises the expectation that there is a relationship between age and size. Furthermore, since firms with higher structural inertia are expected to have greater chances of survival it can be anticipated that the surviving firms also show signs of this inertia, such as fewer changes in product or activity, or from a spatial view fewer changes in location, than 'average' firms. Next to this age-inertia relationship, it can be expected that there is also a relationship between age-size and inertia. This recalls question 4 posed in Chapter 1: "What are the characteristics of old firms, in terms of size, location type, relocation behaviour, products, sector etc.? And do old firms differ from younger firms in these characteristics? Is there a causal relationship between old age and other firm characteristics, in particular concerning the spatial elements?" Furthermore, in Chapter 3, the historic development of industries in the Netherlands was described, and this question as again raised but from the path-dependence perspective: are these differences based on different historical developments and is the survival of individual firms determined by this? Furthermore, in the current chapter, the importance of identity and embeddedness have become clear and these are also incorporated: how important are identity and old firms' embeddedness in their local environment? Considering all of these facets, the overall question becomes: are there relationships concerning firm, age and location in the spatial-economic context?

In the sections above, many questions have been raised about the variables as discussed in Section 4.4.1. From this, the following hypotheses are advanced:

1. Old firms have different characteristics than younger firms.
2. With increased age, firms develop inertia concerning locational behaviour.
3. With increased size, firms develop inertia concerning locational behaviour.
4. Firms founded in adverse environments have less relocation history.
5. Locational behaviour of old firms is first-degree path dependent.
6. Old firms are positively locked-in to their location, product and identity.
7. Old firms experienced adaptation at early times in their lives concerning growth and location, product and market.
8. Old firms' identities are the result of path-dependence.
9. Old firms' generally inert behaviour is linked to the preservation of their identity, in order to maintain high levels of accountability and reliability.

4.5 Concluding remarks

Path dependency creates durable commitments for firms and the outcome depends on relatively minor events during their initial development and location. As time

progresses, a firm's development becomes increasingly important as a determinant of choices regarding new investments in, for example, complementary plants. Accepting this, it is no longer that important whether these structures are the outcome of a deliberate choice, or the outcome of an accidental event. What really matters is that the created inertia, or persistence, of the firm has consequences for the present state of the firm. In this research several specialisation patterns and other causes of importance are observed, including: industry development, specific location, and product. The old firm, and its present state and condition, results from interactions with its environment and caused by historical circumstances, entrepreneurial achievement and locational advantages.

From organizational ecology, we learn that firms survive, on the one hand, by adapting, for instance, process innovation or locational adaptation and, on the other hand, by inertia: the increasing levels of reliability and accountability. Several factors influence this fragile balance between levels of adaptability and inertia within old firms. History matters, since this is the core of path-dependence. Legitimation is important for old firms, in earlier life by creating a market and, nowadays, by maintaining this market. Levels of competition influence this legitimacy. Levels of reliability and accountability are reflected in the identities of the firms, creating inert behaviour by firms in terms of location, product, innovation and culture. Tradition is also important for identity, and tradition is by itself a stimulus for inert behaviour by firms, in a so-called positive lock in. Technology and routines come forward again from this positive lock-in, and create tradition and identity.

A firm cannot be studied without paying attention to its region and the latter's institutions: these interact with the firm, and can create a sense of belonging for a firm. This sense of belonging reflects, in turn on the firm's site and situation. Over time, there is less and less chance of firms leaving their current location. Furthermore, the region and the firms' local environment determine the levels of competition and innovation, spatial proximity to other firms (urban versus rural environment) as well as creating either adverse or encouraging environment.

All these factors influence the inert behaviour of firms. On the one hand, for old firms, this inert behaviour is a necessity, because the firm can no longer 'afford' changes since these would affect its levels of trustworthiness. On the other hand, firms need to adapt in order to react to changing environments. Adaptation in old firms will therefore be externally less 'visible'. The changes will be for example, site adaptation, process-innovation, market extension. Whichever, the changes will not affect the firm's identity and its recognisability to its customers. The old firm has become what it is and where it is by its localisation in time and place; this is path-dependent, through interactions with the competition and regional development. Due to this, the old firm is limited in its behaviour if it wants to continue its 'long-life' or 'success.' In short, old firms show a 'liability to changes'. Due to choices in the past, a positive lock-in of the firm's identity is

created, which reflects inert behaviour in some factors and modest adaptation in others. The firm is bounded by its lock-in; major changes will disrupt this, potentially causing higher mortality rates by resetting the clock!