INCREMENTAL INNOVATION

AND THE UTILITY MODEL

Jyväskylän yliopisto
Kauppakorkeakoulu

Pro gradu -tutkielma

2016

Tuomas Pakarinen

Taloustiede

Ari Hyytinen

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄN YLIOPISTON KAUPPAKORKEAKOULU

Tekijä Tuomas Pakarinen	
Työn nimi Incremental innovation and the utility model	
Oppiaine Taloustiede	Työn laji Pro gradu
Aika 1.8.2016	Sivumäärä 86

Tiivistelmä – Abstract

Tutkielmassa tarkastellaan taloudellisen kasvun tärkeintä komponenttia: teknologiaa. Tarkoituksena on määrittää inkrementaalisen innovaation tarkka sijainti innovaatiotypologiassa ja kirjallisuudesta syntetisoidussa teknologian elinkaari-mallissa. Selviää, että yksittäisen teknologian kehitystä voidaan ajatella jatkumona, jolla on vaikutuksia tutkimusprosessin luonteeseen, ja sitä kautta myös markkinarakenteeseen. Tässä kontekstissa inkrementaaliset innovaatiot ovat innovaatioita kuvitellun artefaktin periferisissä komponenteissa, mikä tarkoittaa käytännössä sitä, että ne yleensä seuraavat dominanttia mallia (dominant design). Nämä kaksi konseptia ovat siis sidoksissa toisiinsa. Teoreettisen kontekstin ja konseptien määrittelemisen jälkeen selvitetään miten innovaatiotypologiaa on mahdollista tutkia empiirisesti. Tutkimuksessa esitetään, että yksi mahdollinen tapa olisi käyttää hyödyllisyysmalia proxynä inkrementaaliselle keksinnölle. Empiirisessä osuudessa tutkitaan hyödyllisyysmallia hyödyntäen inkrementaalisten keksintöjen sijoittumista suomalaisilla toimialoilla, sekä näiden keksintöjen teknologista koostumusta. Selviää että inkrementaalinen keksintö syntyy Suomessa pienessä yrityksessä.

Asiasanat: innovaatio, teknologia, inkrementaalinen, radikaali, hyödyllisyysmalli

Säilytyspaikka: Jyväskylän yliopiston kauppakorkeakoulu[1]

1 I wish to thank JSBE doctoral student Jussi Heikkilä, who acted as an assistant instructor of my dissertation work, for useful comments and discussions.

TABLE OF CONTENTS

1 INTRODUCTION

It can be argued that one of the most fundamental reasons for economic growth can be found in invention of new products and processes. Or expressed differently: societal change can be said to be intertwined with technological change. Yet, what do we economists know about technological change? With a frustrating frequency, discussion around this topic usually collapses to the salience of innovation. What, then, do we know about innovation? Most often this topic is viewed in the contrasting light of dichotomies, such as incremental vs. radical innovation, drastic vs. non-drastic innovation, significant vs. less-significant innovation, and so forth. However, these dichotomies vary in meaning and substance.

Consequently, in the nomenclature that is appointed here and now, the question that will be asked in this study is the following: *What is an incremental innovation*? This facet of the underlying dichotomy is chosen because it is usually simply defined as "not-radical" (Dahlin & Behrens, 2005), which can be seen as an unsatisfying definition. Furthermore, since radical inventions are so rare the antonym to the dichotomy must represent the vast majority of all inventions and subsequent innovations. This means that inventions that can not be classified as radical should constitute the norm in what is technological change. However, here it will be proposed that this simple dichotomy is not sufficient to capture the complexity inherent in the possibilities of improving most products.

Theoretical contexts that touch the topic of incremental innovation include *innovation and technology studies* and *industrial organization literature*. In addition, it will be proposed that the trajectory-like process of incremental innovation can also be thought of as a process of learning, such as learning-by-doing (Arrow, 1962a)[2]. Examination of innovation typology requires the involvement of several lines of literature because the rendition of innovation itself can be said to be incoherent: Innovation can be thought of as the concretion of elements that constitute it (innovation and technology studies) and it can be thought of as the market influence that it commands (IO). This has resulted in the use of disjointed terminology, research methods, and focus amongst innovation related studies. Yet, this shortcoming is exactly what I hope to remedy.

First, innovation and technology studies. Much of these ideas can be represented with the concept of a *generic technical solution*, which can be thought as a triangular bond of a need, a device, and the knowledge that connects the two (Zuscovitch, 1986). Moreover, this triangle can also be thought of as a *technological paradigm* (Dosi, 1982; Dosi & Nelson, 2013). Incremental innovation can be approached from each corner of this triangle. Furthermore, it can also be thought of as normal problem solving activity of which the problems to be solved are demarcated by the prevailing technological paradigm (Dosi, 1982). That is, according to evolutionary economics technological change has a momentary direction, an it will be manifested as a series of incremental innovations.

Second, product life cycle theory first proposed by Abernathy and Utterback (1978) and subsequent works associated with these seminal ideas (Anderson & Tushman, 1990; Henderson & Clark, 1990) . Here, the definition of incremental innovations can particularized by abandoning the strict dichotomy between radical and incremental innovation and allowing innovations to be differentiated by multiple criteria. However, this also complicates the desired definition. Now incremental innovation can be associated with the maturing of an industry: A paradigm initiates a new industry or completely regenerates an old one[3], and depending on the underlying product, it can lead to a *dominant design* (Dosi & Nelson, 2013) which is subsequently improved incrementally inside established firms (Abernathy & Utterback, 1978; Anderson & Tushman, 1990). That is, innovations that come after the formation of a dominant design

2 The view that incremental innovations cen be thought of as a form of learning is elaborated in appendix B.

3 These two things can be thought of as the same thing. However, a technological discontinuity – a paradigm – can obviously be of any specification and size as long as it designates new problem solving directions.

must be incremental ones. However, there is no coercive reason for the formation of a dominant design, and hence, incremental innovations can only apply to certain kinds of industries. Industries that produce artifacts that can be characterized as complex systems (Murmann & Frenken, 2006)

Industrial organization literature, which is part of standard neoclassical economics, examines industrial evolution in terms of incentives for innovation that originate in the surrounding market structure. Here, the watershed that differentiates large innovations from small ones is *drasticity* (Arrow, 1962b). A drastic innovation is one that constitutes a large enough cost reduction or quality increase that is sufficient to create a monopolist out of the innovator. In addition to this definition of innovation size, game theoretic models shed light on the motives of individual agents. The well known debate of Gilbert and Newbery (1982, 1984) and Reinganum (1983, 1984a, 1985) formulate a *reason* for non-drastic, potential incremental, innovation: established firms want to keep their position in the market, and this happens by disincentivizing challengers from creating marginally superior products and processes. Patents, as a right to exclude others from producing the patented invention, are essential for this kind of behavior. Gilbert and Newbery (1982) even propose that established firms might invent and patent prominent products and not produce them.

Altogether, incremental innovation seems to have a time, a place, and a reason behind it. Some inventions are entirely new to world. These, novel and potentially radical inventions encompass entirely new *knowledge* which manifests itself in a new base principle, which proceeds from an effect that is something usable and exploitable (more in Arthur, 2007). Subsequent inventions that utilize this same base-principle can be called incremental, even though a more appropriate name would simply be *subsequent* invention. This is because these subsequent inventions can be of interminable size and shape. Further definition of an incremental innovation requires another concept. A *dominant design* can emerge for various reasons (more in Murmann & Frenken, 2006) and it removes e.g. architectural innovations (more in Henderson & Clark, 1990). It can be thought of as the second guide post (Sahal, 1985) in industrial evolution. Innovations that come after this watershed moment are surely incremental. All in all, these definitions will situate incremental innovations inside established and scale-intensive firms. Lastly, IO literature can explain the behavior of the firms at this mature stage: These innovations that are motivated by fear of losing out to competitors or new entrants. T*hey are there to keep the status quo.* More in chapter 2.4.

How can such innovations be identified? Theory is just a conjecture without data. One such mean is the *utility model* (Beneito, 2006). Essentially the question is this: *Are innovations protected by utility models good proxies for incremental innovations*? If they are, utility models would provide a method for testing the various theories surrounding technological change and innovation size. The incrementality of Finnish utility models will be analyzed by assigning these 'petty patents' to different technologies designated by IPCs, as well as, by connecting utility models to European industry classifications (NACEs). Finland represents a manifold case for the aforementioned study since it is one of the most complex (8^{th} to be exact) economies in the world according to Observatory for Economic Complexity (OEC, online cition).

It will be found that utility models are relatively more important in certain kind of industries that can be classified as *supplier dominated* (Pavitt, 1984). However, It will also be found that nearly all industries that patent also use utility models. The more scientific the industry is, e.g. electronics, the less emphasis it puts on utility models. Moreover, the more firms an industry has and the lower the average turnover of this industry, the higher the number of utility model IPC connections will be. Small firm use utility models because they are cheap compared with patents. This also explains the low technological content correlations with patents.

This dissertation is organized as follows: Section 2 will establish a definition for incremental innovation in a theoretical context: Chapter 2.1 will review the essential concepts for understanding innovation related studies, chapter 2.2 will establish a picture of what is meant by incremental innovation in theoretical innovation and technology literature, chapter 2.3 will review what is meant by a similar concept – a drastic innovation – in industrial organization literature, and chapter 2.4 will summarize and define what an incremental innovation could be. Section 3 will review how “small” innovations are studied in empirical innovation studies and discuss how such methods relate to the study of incremental innovations. Section 4 will research the applicability of utility models as proxies for incremental innovations. Section 5 will conclude and discuss empirical results, as well as, the proposed theoretical concept. Appendixes A, B, and C will further the discussion.

2 REVIEW OF THE PRIOR THEORETICAL LITERATURE

2.1 Semantics of innovation typology

First it is necessary to reiterate what is meant by 'innovation'. Greenhalgh and Rogers (2010) define innovation as the *application* of new ideas to the products, processes, or other aspects of the activities of a firm that lead to increased value added, or benefits to consumers, or to other firms. This description hints at a separation between invention, innovation, and diffusion of innovation – segregation that dates back to Schumpeter (1942). Nonetheless, most researchers do not distinguish between invention and innovation (Carlino & Kerr, 2014). An invention only becomes an innovation after the commercialization of the underlying product or process (e.g. Schumpeter, 1939; Greenhalgh & Rogers, 2010). Thus, inventions do not necessarily lead to innovations, albeit innovations require inventions, blueprints, plans, and most importantly, *investment* (Greenhalgh & Rogers, 2010).

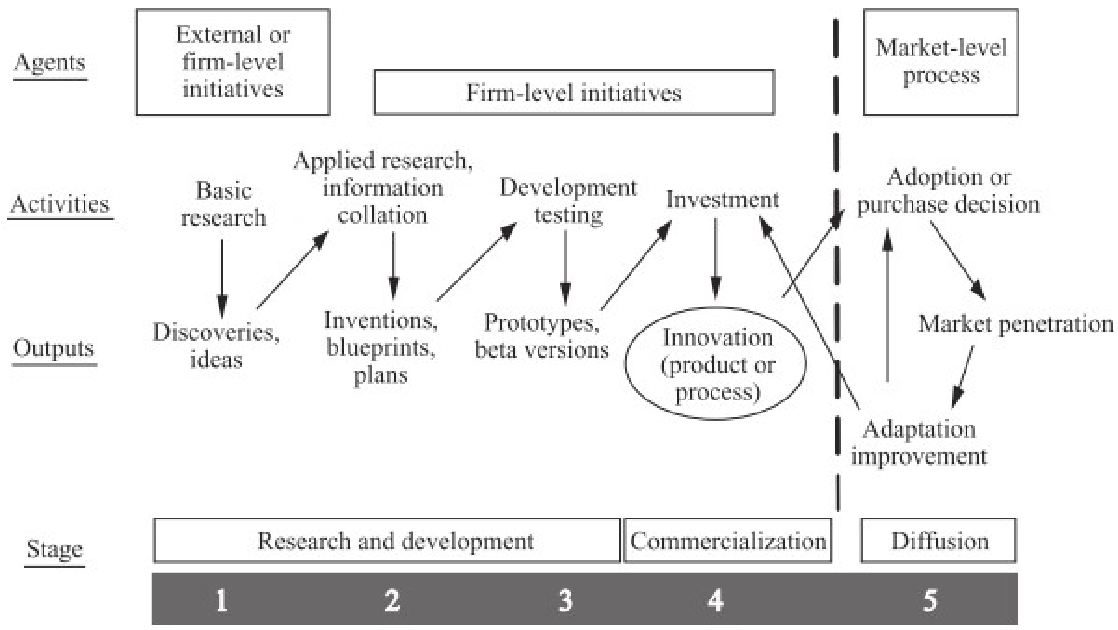
Furthermore, innovation and invention are not separate, monotone concepts. There are distinct invention and innovation types. The most commonly used types are three sets of dichotomies (Carlino & Kerr, 2014). The first distinction is between *product* and *process* invention (innovation). Product invention refers to the introduction of a new product or to an improvement in an existing product, while process invention refers to the introduction of a new process for making or delivering goods and services (Greenhalgh & Rogers, 2010). However, this separation is not as absolute as it might first seem; a product of one firm might be useful for the process of another firm, i.e. the production and the use of chain saws. The second distinction, which is related to product-process separation, concerns the origin of an innovation, i.e. whether the innovation is *external* or *internal* vis-à-vis the organization of a firm (Carlino & Kerr, 2014). The third distinction, also related to the product-process separation, is between *tangible* and *intangible* inventions (Greenhalgh & Rogers, 2010).

Inventions and innovations can also be viewed in terms of their antecedents and their effects. An invention can be viewed in terms of its ex ante characteristics or it can be viewed in terms of its ex post technological impact (Verhoeven, Bakker & Veugelers, 2016). Characteristics relate to the way an invention differs from other inventions in the way that it incorporates new knowledge or moves away from prevailing practices, and technological impact relates to the way an invention affects the development of future inventions (Verhoeven, Bakker & Veugelers, 2016). In a similar manner, innovations can also be divided into ex ante characteristics of an innovation – the invention – and its ex post market impact. Hence, there are three dimensions to an innovation: underlying invention, its technological impact, and its market impact.

As was mentioned in the introductory chapter, the terminology that is the essence of this study encompasses perhaps the oldest invention related dichotomy: one between *incremental* and *radical* innovation. Incremental innovation, first studied by Usher (1929), refers to "minor" cumulative improvements to existing products. I will return to this topic time and time again throughout this study. Radical innovations, on the other hand, are also known as disruptive (e.g. Anderson & Tushman, 1990), and they might be drastic (Arrow, 1962b). Innovation's drasticity relates to its effects on the prevailing market structure. Nevertheless, it should be noted that the terminology surrounding radical and incremental innovations has not solidified to a coherent whole (Dahlin & Behrens, 2005). Therefore, this study hopes to further this goal of achieving a concordant whole.

Figure 2.1 illustrates all that is essentially part of the *innovation process*, where innovation itself is the end result of this process. The vertical axis denotes actions and outcomes of agents – firm, non-firm, and market level processes. Greenhalgh and Rogers (2010) segregate innovation process into five stages (also the horizontal axis): The process starts from basic research which could also be called human knowledge, i.e. something that enables ideas that can be economically beneficial. These matters are largely determined by factors external to the market processes. The second stage encompasses applied research whose main purpose is to create novel inventions. The third and fourth stages transform an invention to an innovation. These matters are largely carried out within a firm. However, innovation is not the absolute end result of the innovative process which is continuous. Market forces produce a feedback mechanism: In order to keep their position in the marketplace, innovators must make sure theirs' is the most preferred variant of the invention.

Figure 2.1: What is meant by innovation?



Adopted from Greenhalgh and Rogers (2010

However, even this picture offers a simplistic view of innovation and, hence, it is not possible to build a definition of an incremental innovation on such loose semantic analysis. The task at hand requires understanding technological evolution itself. This is the topic of next chapter. What should be discernible from this semantic analysis is that invention and innovation are very different things, inventions themselves can take a myriad of forms, and both inventions and innovations should be thought of as their antecedents and their effects on the surrounding technology and market structure.

2.2 Technological evolution and incremental innovation

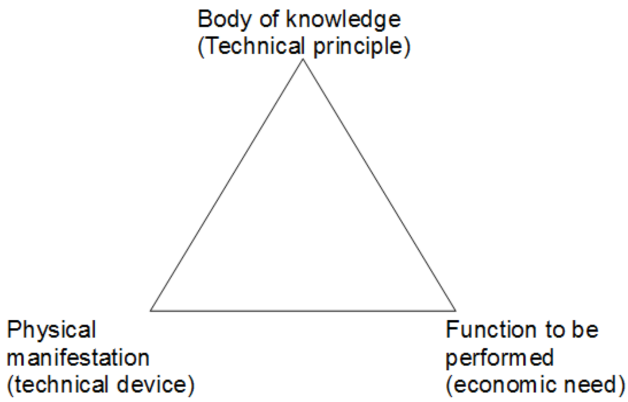
The process of technological change can be understood as an evolutionary process where the search of profitable opportunities is refined and guided by selection mechanisms originating in human needs and market competition (Dosi & Nelson, 2013). However, this statement should not be mistaken as a direct analogy with biological evolution; what is evolving here is much more than a set of genes, and hence, unlike biological mutation, technological variation is not totally random. However, this purposefulness of search does not mean that inventors, innovators, or investors can ex ante and with full certainty know the impact that an individual invention will have on the future technological development (Dosi & Nelson, 2013). When the technological variation that this uncertainty brings about is combined with the aforementioned selection mechanisms, the end result is a process that resembles evolution.

In order to understand how technological evolution works and what could be meant by incremental innovation, one must first establish a picture of what is meant by technology. Some of the earliest economic theories to tackle technological change can be characterized either as 'demand-pull' and 'technology-push' theories (Nelson & Winter, 1977; Dosi, 1982; Pavitt, 1984). However, as Mowery and Rosenberg (1979) point out, both of these factors are necessary for innovation, indicating that the nature of the phenomenon is indeed cyclical and cumulative. In this study, I will therefore employ Zuscovitch's (1986) idea of a generic technological solution which combines demand-pull and technology-push aspects of technological change. According to Zuscovitch (1986):

> "a *specific technology would appear to be a theoretical unit which gives structure to elementary small innovations according to a particular cumulative path*" (p. 175).

He reiterates this definition of a generic technological solution as having three constituent parts: a body of knowledge (technical principle), a physical manifestation (technical manifestation), and a function to be performed (economic need) (Zuscovitch, 1986). These three parts are illustrated in figure 2.2 as corners of an triangle which represents the generic technology. It can also be thought that the triangle consists of one act: a man (top angle) putting a device or a thing (left angle) to a purpose (right angle).

Figure 2.2: A generic technology



Adopted from Zuscovitch (1986)

A generic technology is therefore means to fulfill a human purpose that is build around an idea or a concept that links an effect with a need. Yet, technologies are also combined from smaller sub-systems and eventually, from components, which together form the *hierarchy* of the technology and the *architecture* of the device. This hierarchy needs to be build around some central idea of the technology: the *base principle* of operation. This base principle is what connects an effect the with the need. For example, the base principle of a jet engine is to compress and then heat air in order to create thrust. (Arthur, 2007). The triangle of figure 2.2 encompasses these matters. Base principle is represented in the top corner and device matters in the lower left corner.

Another way to view the triangle of figure 2.2 – a generic technology – is to think of it as a *technological paradigm* (Dosi & Nelson, 2013). According to Dosi (1982) technological problem solving resembles scientific problem solving: Modeled after Kuhn's (1962) scientific paradigms, technological paradigms[4] relate to the cognitive frames of technological professionals that direct their attention to specific problems and solutions (Dosi & Nelson, 2013). However, the analogy with science should not be taken as an identity for much of technological knowledge is tacit nature, i.e. it is a product of experience (Dosi, 1982). A technological paradigm can be defined approximately as an "outlook"

4 It is important to notice that these paradigms are *micro-technological*, meaning that they concern particular technologies (e.g. semiconductors, light bulbs, or cars), not macro-technologies such as 'flight' or 'computing' (Dosi & Nelson, 2013).

that defines the relevant problems and the methods for addressing them (Dosi, 1982). This outlook comprises of "positive heuristics" and "negative heuristics" that embody strong prescriptions of the directions of technical change to follow and to avoid (Dosi, 1982).

According to Nelson and Winter (1977) both the individual R&D project and the choice of R&D variant are *heuristic search processes:* "a heuristic search process is an activity that has a goal, and a set of procedures for identifying, screening, and homing in on promising ways to get there" (p. 52). A set of heuristics regarding R&D project selection can be in turn regarded as an *R&D strategy* (Nelson & Winter, 1977). Following demand-pull and technology-push theories, R&D strategy can be based on either capabilities, perceived demand characteristics, or both. Each of these strategies will have associated payoffs and risks that will most likely differ greatly from each other, demand-pull being the more common one and the one with lower expected payoff (Nelson & Winter, 1977). Nelson and Winter (1977) refer to a continuum of these favorable strategies as *natural trajectories* of technological change. They also identify two of these natural trajectories: exploitation of latent scale economies and increased mechanization of operations that have been done by hand. In the 21st century one additional natural trajectory could arguably be the automatisation of these mechanized devices.

Following Nelson and Winter (1977), Dosi (1982) outlines a *technological trajectory* as "the pattern of 'normal' problem solving activity (i.e. of "progress") on the ground of a technological paradigm" (p. 152). Therefore, a paradigm sets boundaries for any future developments and designates the direction to be followed. Or in the terminology of Nelson and Winter (1977), it constrains the list of possible R&D strategies. Hence, with the introduction of time, the technology will progress along the path pointed out by the paradigm. These factors together form the technological trajectory – *ex post*. Two effects of trajectories stand out: First, a trajectory orders and confines the generation of new or modified products and processes but does not remove generation of variety (Dosi & Nelson, 2013). Second, trajectories constrain the uncertainty inherent to innovation[5] (Dosi & Nelson, 2013). Nevertheless, establishment of a

5 It can be argued that technological change (invention, innovation, and diffusion) involves major uncertainty, mostly the Knightian kind. Individual "inventors" do not know where the inventive process is heading since the process consists of thousands of other inventors. The same applies to innovation process. Furthermore, the diffusion of innovations is by itself uncertain. Historical evidence points to the fact that novel

technological trajectory is not nearly enough to create unbiased technological expectations or to 'probabilize risk' (Dosi & Nelson, 2013).

What does innovation typology mean in this context? A rough descriptions of qualitative differences in innovation can be achieved with the aforementioned concepts – technological paradigm and trajectory. As Dosi (1982) put it:

> "*...incremental innovation versus radical innovations can be reinterpreted in terms of 'normal' technical progress as opposed to new emerging technological paradigms*" (p. 158).

Now that what is meant by a generic technology is specified and the cognitive frames of R&D professionals have been mapped it is time to study the first part of technological evolution: invention. According to Arthur (2007):

> "*...invention is a process of linking some purpose or need with an effect that can be exploited to satisfy it. It proceeds from a need for which existing methods are not satisfactory, which forces the seeking of a new principle (the idea of an effect in action); or from a phenomenon or effect itself – usually a freshly discovered one – for which some associated principle of use suggests itself. Either way, translating this principle into physical reality requires the creation – and combination – of suitable working parts and supporting technologies*" (p. 275).

Hence, there are two ways through which novel technologies can arise: Firstly, inventions can be motivated by a need that can concern any part of human existence. Here the old saying saying holds true: *necessity is the mother of invention*. Initial advances in mobile phone technologies, for example, were motivated by military needs. Secondly, inventions can also come about through observation of a novel phenomenon. Fractal shaped antennas, for example, were simply observed to be more efficient than any other shape, and hence, they are now in every mobile phone.

Arthur's (2007) view of inventive process is one of recursive (starting from higher levels in the hierarchy and moving to lower ones) problem solving. Yet, everything revolves around the base principle. Hence, Arthur (2007) defines a radically novel technology as one that satisfies a need with a novel base principle. Conversely, inventions within the bounds of the old technology (i.e.

---

innovations are not utilized in their original form, but require a cascade of improvements before they can be diffused (Rosenberg, 1976).

inventions that use the same base principle) can be classified as incremental improvements.

Arthur's (2007) idea of a base principle that designates the route for recursive problem solving activities involves a remarkable similarity with Dosi's (1982) idea of a technological paradigm. However, they are not the same thing. As Nelson and Winter (1977) pointed out in their seminal article, both R&D project and R&D project choice need to be thought of as heuristic search processes. In this view, technological paradigms guide the R&D project choice and a base principle guides an individual project, at least when the inventive search process is initiated by a need.

Next it is time to describe 'small' innovations with the help of the lower left corner of Zuscovitch's (1986) triangulum of a generic technology: the physical device. Murmann and Frenken (2006) define technology as a man-made system that is constructed from components that function collectively to produce a number of functions for users. Thus, technology can be seen as a *complex system* (Murmann & Frenken, 2006). Small changes in some part of a complex system can change the collective behavior of the whole system. In this view, incremental trajectory-like[6] nature of technological development results from the interdependencies between parts of a complex artifact – the technical product (Murmann & Frenken, 2006). Arthur (2007) called this recursiveness of the inventive process, albeit he did not define a technology as a complex system but a nested hierarchy.

Arthur's (2007) analysis of technology included four parts: base principle, central assembly, other sub-systems, and architecture. Similarly, the complex system view of technology that is proposed by Murmann and Frenken (2006) also includes four parts: operational principle, core, periphery, and design space. These two views of technology are identical in all respects except for design aspects: The complex system view of technology enables reversiveness in recursiveness. That is, Arthur's (2007) view of major inventions meant a top-down approach to hierarchical problem solving activity, where (radically) new inventions start from a novel base-principle that is either freshly discovered or which is combined with a specific need. The complex system view of technology, on the other hand, enables a bottom-up approach to inventions.

The view that technology can be seen as a complex system consists of several factors. Firstly, operational principle relates to the knowledge that is needed to build a device that works in a particular way and it is analogous to

6 Also described as path-dependency of innovation .

Arthur's (2007) base principle. Secondly, complex hierarchical systems are not directly made of elementary particles but of subsystems, and eventually, individual components – the elementary particles. *Interdependencies* between the components imply that some parts of a device can not be improved without making accompanying inventions in other parts. Thus, when an artifact becomes more complex, in the sense of proliferance of interdependencies, a successful innovation becomes more difficult. Relatedly, interdependencies can also lead to a diminished role of prices in factor substitution. Secondly, in order to be a functional whole, the elementary particles (components) need to be arranged *hierarchically*. These hierarchies can be either hierarchies of inclusion or hierarchies of control. According to hierarchy of inclusion each subsystem is free to go through its own technology cycle[7], while hierarchy of control implies that some components control other components within a given subsystem. Thirdly, components can be classified as either core or peripheral with their *pleiotropic*[8] characteristics. Designers create the technical characteristics of an artifact while customers decide on the basis of artifact's service characteristics. (Murmann & Frenken, 2006).

Murmann and Frenken (2006) thus define pleitropy of a technical component as:

> "*...the number of functions affected by this component—meaning the number of service characteristics that will change their value when this component in the system is changed*"(p. 941).

Consequently, high-pleiotropy components are defined as core components and low-pleiotropy components as peripheral components. Once core components have been selected further advances happen in peripheral components. However, even though changes in core components are unlikely to succeed, they are still possible. Yet, they do require changes in other components and maybe even in the way components are put together – the system architecture. It is because of this that improvements in core components may initially lead to

7 I will return to this concept after I have discussed what a dominant design is and how it relates to incremental innovations.

8 In biology pleiotropy refers to the number of attributes affected by a particular gene in the genotype. Murmann and Frenken (2006) use this concept to study technological characteristics of complex artifacts. In their view "technical characteristics (the components) make up the "genotype" of a product technology, and service characteristics (the product attributes) make up the "phenotype" of a technology" (Murmann & Frenken, 2006. p. 940).

poorer aggregate performance. Hence, it can be stipulated that invention's incrementality increases when it applies to lower levels of the system hierarchy. Or when expressed conversely:

> *"Moving upward in the systems hierarchy increases the radicalness of innovation in terms of the scope of new knowledge required*" (Murmann & Frenken, 2006. p. 944).

These ideas can be contrasted with earlier innovation typology related studies. According to Henderson and Clark (1990) innovation can be typified according to its effect on *core design concepts* and, on the other hand, according to its effects on the *system architecture* (Henderson & Clark, 1990). In this view innovation encompasses two kind of knowledge:

> "*First, it requires component knowledge, or knowledge about each of the core design concepts and the way in which they are implemented in a particular component. Second, it requires architectural knowledge or knowledge about the ways in which the components are integrated and linked together into a coherent whole*" (Henderson & Clark, 1990. p. 11).

Table 2.1 extrapolates the position of incremental innovation in this division. Incremental innovation can be seen as the opposite to radical innovation in two dimensions: core concepts (base principle or operational principle) and linkages between core concepts and components (pleiotropy relations of components) (Murmann & Frenken, 2006).

Table 2.1: four kinds of innovations Core (design) concepts

Linkages between core concepts and components (system architecture)	Reinforced	Overturned
Unchanged	Incremental innovation	Modular innovation
Changed	Architectural innovation	Radical innovation

Adopted from Henderson and Clark, 1990

Two other forms of innovations emerge from this distinction: a modular innovation and and an "architectural" innovation. A modular innovation can be separated from an incremental innovation in that it *changes* the core concept of a design leaving the system architecture untouched (Henderson & Clark, 1990).

Architectural innovation, on the other hand, is "...the reconfiguration of an established system to link together existing components in a new way" (Henderson & Clark, 1990. p. 12). Nonetheless, it is worth noticing that the distinction presented above is not absolute – An invention can be *more* radical, architectural, incremental or modular (Henderson & Clark, 1990). Altogether it can be stipulated that i*ncremental innovations reinforce the established design* (Henderson & Clark, 1990).

Yet, what is an established design? The complex system view of technology offers a definitive answer to this question: a *dominant design*. The standardization and solidification of higher pleiotropy (core) components equals the emergement of a dominant design (Murmann & Frenken, 2006). After these core components and linkages between them have been solidified, the problem solving activity is bound to move to lower pleiotropy components (Murmann & Frenken, 2006). This is, because a change in core components needs to be accompanied with changes in lower pleiotropy components, in turn, this is because of the nature of the pleiotropy concept it self (Murmann & Frenken, 2006). According to Murmann and Frenken (2006):

> "*This explains why changes in core components of a technology take so much time and effort before they are successfully introduced. It also explains why, at any level of an artifact hierarchy, technology cycles are triggered by the substitution of core components*" (p. 942).

In the light of what was imparted above and earlier, innovation typology and complex system view can be linked the following way. Innovation can be determined radical in multiple ways, either in new knowledge required (i.e. invention's antecedent), performance enhanced (i.e. invention's technological ramifications) (Murmann & Frenken, 2006), or in competitive consequences. This is illustrated in table 2.2, which separates radicalness into two dimensions. Inventions that are radical in both dimensions are surely radical. In the system hierarchy, innovation becomes more radical in knowledge dimension as it happens higher in the system hierarchy, i.e. the new design applies a new principle to more and more individual components (Murmann & Frenken, 2006). The same, on the other hand, does not apply to performance dimension. Innovations that take place at lower levels of the hierarchy *can* have more radical effects than innovations that involve the entire system (Murmann & Frenken, 2006). This is because changes in higher pleiotropy components requires additional changes in supporting technologies, i.e. low pleiotropy components, and all this takes *time* (Murmann & Frenken, 2006).

Altogether, what is crucial for the definition of an incremental innovation is the emergement of a dominant design. It solidifies high pleiotropy (core) componenents, forcing the problem solving to move to lower pleiotropy (peripheral) components in the system hierarchy (Murmann & Frenken, 2006). How, then, does a dominant design come to existence, i.e. "emerge"?

Table 2.2: Other four kinds of innovations Performance improvement

Scope of new knowledge	Low	High
Small	Incremental innovation	Radical innovation sense 1
Large	Radical innovation sense 2	Radical-square (r^2) innovation

Adopted from Murmann and Frenken (2006)

First concocted by Abernathy and Utterback (1978), the concept of dominant design has found a place in the innovation and technology research. According to Abernathy and Utterback (1978) development of an industry, from start to maturity, is characterized by change in the behavior of firms. In the beginning firms are small, entrepreneurial, and focus on economies of scope, while in the maturity firms are large, established, and focus on economies of scale (Abernathy & Utterback, 1978). According to Abernathy and Utterback (1978) in the mature stage innovation is typically incremental and has a gradual, cumulative effect on productivity. The incentive for this kind of innovation is usually *cost reduction*, but it does not mean that there are no increases in the performance (Abernathy & Utterback, 1978). On the other hand, newly born industries are characterized by 'radical'[9] innovation, which is focused on achieving better *performance* and meeting *customer demands* (Abernathy & Utterback, 1978).

A dominant design may emerge with in a product category for a number of reasons: Murmann and Frenken (2006) identified five causal mechanisms that might lead to its emergement. Firstly, a dominant design might be a compromise among different functional characteristics of a device (Abernathy & Utterback, 1978; Utterback & Suarez, 1993). In this view, a dominant design is

9 Henderson and Clark (1990) would call these innovations 'modular' or 'architectural'.

a deterministic milestone along device's development path. Secondly, a dominant design enables economies of scale (Klepper, 1997). In this view, a dominant design is a consequence of first-mover advantage. Thirdly, network externalities[10] may force a certain design to be adopted. In this view too, a dominant design is a consequence of first-mover advantage. The other two causal mechanism included possibilities that an initial lead in market share coupled with some kind of self-reinforcing process might lead to a wide adoption of a particular design, and the possibility that other sociological, political, and organizational dynamics might make it less probable that a dominant design is selected through a market process. (Murmann & Frenken, 2006).

Other scholars have studied how the emergement of a dominant design will affect the evolution of an industry (e.g. Anderson & Tushman, 1990). In the technological life cycle (TLC) model the formation of a dominant design transforms the form of competition from "Schumpeterian stage" to the stage of oligopolistic maturity[11] (Anderson & Tushman, 1990; Klepper, 1997). Furthermore, the emergence of a dominant design changes the focus of R&D efforts from product to process innovation (Klepper, 1997) and these conditions might be more suitable for larger firms (Utterback & Suarez, 1993). Altogether, the emergence of a dominant design can be seen as a watershed moment in the evolution of a product class, and hence, the industry. Abernathy and Utterback (1978) see that a dominant design emerges only once in the evolution of a product class. Conversely, Anderson and Tushman (1990) see the emergence of a dominant design as a recurring theme in the evolution of a given product class.

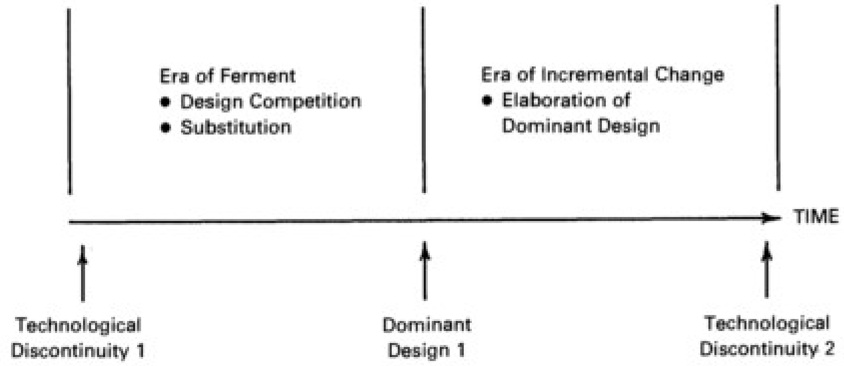
Anderson and Tushman (1990) propose the following dynamism: A technological discontinuity (i.e. an invention that connects a new base principle to a need), which can be interpreted as a radical innovation, initiates an era of intense technical variation and selection, which culminates to the formation of a dominant design. In the complex system view this means the solidification of core components. After this initial stage of competition, a period of incremental

10 A network effect (also called network externality or demand-side economies of scale) is the effect that one user of a good or service has on the value of that product to other people

11 Dosi (1982) also thought that advancement along the technological trajectory would lead to a change in the competitive landscape. Later dominant design was incorporated into paradigmatic literature (Dosi & Nelson, 2013)

technological change follows, which may again be broken by a subsequent technological discontinuity. Innovation is incremental in the sense that it reinforces the established design, and this design might be appropriated by a single firm[12]. Anderson and Tushman (1990) choose to call the two temporal periods that follow from the previously described logic as "era of ferment" and "era of incremental change". Figure 2.3 illustrates the idea.

Figure 2.3: A technology life cycle



Adopted from Anderson and Tushman (1990)

Initially, an era of ferment is ushered in by an invention that in the beginning is crude and experimental. This era is characterized by substantial product-class variation and experimentation, as well as, two different kinds of competition: competition between technological regimes and competition within the new regime. This means essentially that several versions of the radically new innovation appear. The era of ferment is ended by the emergence of a dominant design. The emergence of a dominant design begins an era of incremental change. At this stage product variation evolves around the standard established by the dominant design and not in challenging it. Here, incremental innovation is the driving force in lowering the costs of production and in minor quality improvements. Thus, the competitive landscape is distinct from that of the era of fermentation:

12 The existence of patents as forms to exclude others from producing an invention – or a design – surely make this possibility reasonable. However, the most important fact is the patenting of the design of core components, especially if a particular design of these core components results in superior performance.

> "*...new designs must win market share from an entrenched standard that is well understood within the marketplace, whose costs have been driven down an experience curve, and which often benefits from centrality in a network of supporting technologies*" (p. 617). (Anderson & Tushman, 1990).

Murmann and Frenken (2006) extend these ideas to every hierarchical level of the complex system hierarchy. According to them products that are based on complex systems evolve in the form of a nested hierarchy of technology cycles (Murmann & Frenken, 2006). This means that a process that was depicted in figure 2.3 applies to all (core, peripheral, and component) levels of the system hierarchy. This view have several implications which I will discuss further in the concluding chapter (ch. 5).

Nevertheless, the emergence of a dominant design is by no means a deterministic inevitability in every industry. For example, there is no sign of dominant designs in pharmaceutical technologies (Dosi & Nelson, 2013). Hence, it must obviously be stipulated that dominant designs can only exist within product categories that have a design, i.e. a base principle and *a hierarchy of supporting technologies and components*. If the product consists of a base principle and little else, an improvement to the function that it performs always involve a novel base principle, and thus, constitute a novel invention as was characterized by Arthur (2007). Hence, incremental inventions and innovations only apply to certain kinds of technologies and industries.

Obviously every industry has its own peculiarities. On the other hand, it is possible to find some consistencies between different sectors of the economy, and hence, approximate the "location" of incremental innovations amongst industry types. In the search for possibilities firms are not indifferent about the choice of technology that they choose to pursue (Pavitt, 1984). They research what they do and know about. Thus it can be expected that different principal activities generate different technological trajectories (Pavitt, 1984). Pavitt (1984) used these ideas to compartmentalize all firms into four categories. His taxonomy includes four kinds of sectors:

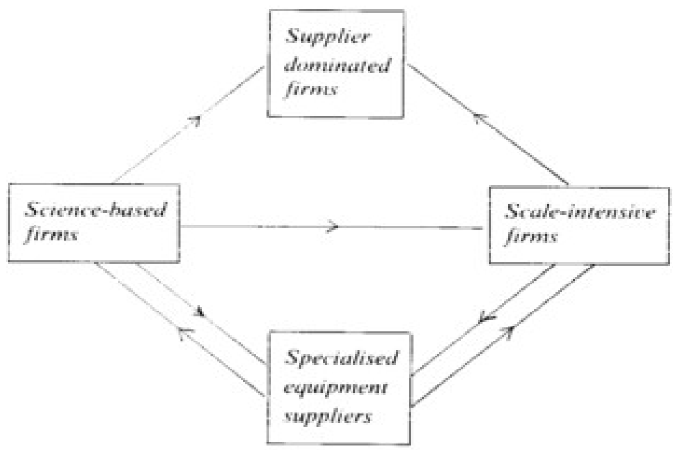
i) *Supplier dominated*: Industries that mainly receive new technologies from suppliers of machinery and intermediate inputs, e.g. textiles, lumber, wood products, and metal products.
ii) *Specialized suppliers*: Industries that produce industrial machinery and equipment.
iii) *Scale intensive sectors*: Industries that rely on scale economics and whose

"...innovative opportunities partly endogenously generated and partly stemming form science based inputs" (Dosi & Nelson, 2013. p. 24)[13].

iv) *Science based industries*: Industries that rely on scientific developments for innovative opportunities. E.g. pharmaceuticals, microelectronics, chemistry, and bio-engineering (Dosi & Nelson, 2013).

These sectors form co-action relationship that is pictured in figure 2.3. Arrows in the picture indicate technology and knowledge flows. That is, some (science based, specialized equipment suppliers, and scale-intensive) firms produce their own technology and also transfer it to other sectors, while others (supplier dominated) firms only receive technology from other sectors. A small construction company, for example, uses tools from scale-intensive firms, who use machines to make them, and these machines in turn have been acquired from specialized equipment supplier firms, who might have included robotics and algorithms in these production machines, which originate in science-based firms. Algorithms and robotics can also be sold directly to supplier-dominated and scale-intensive firms, et cetera.

Figure 2.3: A mapping of inter-sectoral knowledge flows



Adopted from (Pavitt, 1984)

According to Pavitt's (1984) taxonomy and theory, novel scientific ideas and subsequent inventions pass though certain kinds of industries and firms

13 Two further scale intensive sectors can be distinguished: 'discontinuous' complex-product industries such as automobiles and 'continuous' flow industries such as oil refining or steel production (Dosi & Nelson, 2013).

contained within (aptly termed science based firms). These firms then diffuse these ideas to the rest of the economy. Other sectors of the economy rely on *technological developments* for their inventive opportunities, and these opportunities are largely combinatorial in nature (Arthur, 2007). While specialized equipment suppliers operate in a symbiotic relationship with science-based and scale-intensive firms, supplier-dominated firms use mature technologies from other sectors.

The sectors that are more inclined towards incremental innovation are the following: specialized equipment suppliers, scale-intensive, and supplier-dominated ones. However, as incremental innovations take place mainly after the formation of a dominant design (Anderson & Tushman, 1990; Murmann & Frenken, 2006) and it is this design that also enables economies of scale (Klepper, 1997), it can stipulated that incremental innovations can be associated with discontinuous scale-intensive and specialized equipment supplier firms. Moreover, Pavitt (1984) describes supplier-dominated industries as not particularly innovative, as it is mostly comprised of small companies in traditional branches of industry.

To summarize, technological evolution seems to have a structure of its own, one that is partly economic and partly originates in technology itself. As with any creative endeavour, everything starts with an idea – one that connects a perceived need with with an effect that satisfies it, or vise versa, with a novel effect for which a need suggests itself (Arthur, 2007). The idea then needs to be converted to a working device – an invention: This requires experimenting with different kinds of configurations in a device that can be thought of as a hierarchical complex system (Murmann & Frenken, 2006). The problem solving activity starts with a goal of finding a suitable design of essential core components. This era is characterized by substantial product-class variation and experimentation with different kinds of designs (Tushman & Anderson, 1986; Anderson & Tushman, 1990). Eventually, a dominant design might emerge from this design experimentation. A dominant design is defined by the solidification of core components to a specific arrangement (Murmann & Frenken, 2006). The solidification of core components requires that the problem solving activity moves to the peripheral components – this is what incremental innovation is.

2.3 Industrial organization and incremental innovation

Innovation studies that can be associated with industrial organization studies have posited a particular innovation or a sequence of innovations as the starting point in their analysis (Reinganum, 1989). This means that it is no longer correct to speak of inventions' technological antecedents and ramifications. Conversely, one is required to speak of market structure (Tirole, 1988), first-mover advantage, and uncertainty (Reinganum, 1989) as deciding factors in the intra-industrial allocation of *R&D expenditure* and the *identity of the innovating firm*. The initial market structure can be examined either with symmetric or asymmetric models (Reinganum, 1989).

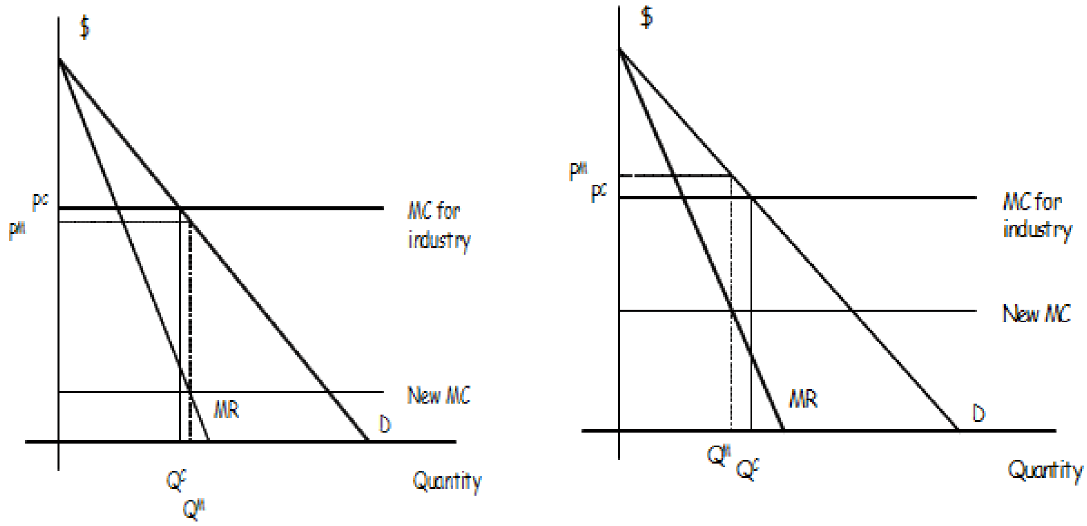
A typical outcome of symmetric (agents start with the same market power) non-cooperative model equilibria is that when compared with models that include cooperative or surplus-maximizing agents, it is associated with too high aggregate R&D spending and too many firms investing in R&D. Hence, the situation is analogous to the problem of commons. Asymmetric models, on the other hand, analyze situations with differing initial market power, anticipated innovation, and technological advantage. The results of asymmetric innovation analysis are indicative of the following:

> "*Results in this area seem particularly sensitive to the presence or absence of technological uncertainty in the production of the innovation. When innovation is uncertain, a firm which currently enjoys a large market share will invest at a lower rate than a potential entrant, for an innovation which promises the winner a large share of the market. When innovation is deterministic, the opposite is true*" (p. 851). (Reinganum, 1989).

However, before these analytical dynamics can be linked with what was imparted in the last chapter, some important concepts must be introduced, such as replacement effect, efficiency effect , and drastic innovation. In his seminal article, Arrow (1962b) argued that the incentive for innovation was greater under competitive conditions than under monopolistic ones. This was later coined as the *replacement effect,* which states that by innovating, the monopolist merely hopes to replace its position in the marketplace (Tirole, 1988). Yet, this effect needs to be contrasted with another effect which states that in a homogeneous-good industry a monopolist will not profit less than two non-colluding duopolists. i.e. the *efficiency effect* (Tirole, 1988).

Furthermore, Arrow (1962b) defined a *drastic innovation* as an innovation that represents such a drastic reduction in production costs that the innovator essentially becomes a monopolist. Figure 2.4 illustrates Arrow's ideas. The vertical axis represents product's price and the horizontal axis represents its quantity. In this most simplistic illustration, all concepts are assumed to be linear. If one assumes perfect competition, pre-innovation marginal cost equals the price. A drastic innovation (left hand side) lowers the marginal cost (in this example also the unit cost) of production to such level that marginal cost equals marginal revenue. A non-drastic innovation (right hand side) doesn't constitute such a drastic cost reduction.

Figure 2.4: A drastic and a non-drastic innovation



Expressed differently, if c are the unit costs before innovation, c' are the unit costs after the innovation, and p is the monopoly price of innovation, then a drastic innovation is one for which p(c') ≤ c and a non-drastic innovation is one for which p(c') > c. This means essentially that the old technology remains a viable substitute for the new one. Moreover, following Arrow's (1962b) argument for process innovations, a product innovation can also be considered either drastic or non-drastic, depending on whether the old technology becomes entirely obsolete and its demand falls to zero upon the introduction of a new product even when this product is introduced at its monopoly price (Henderson, 1993). If this happens, the product innovation can be considered as a drastic one.

However, these concepts all by themselves do not tell much about incentives for different kinds of innovations as indicated by innovation typology: Arrow (1962b) argued that both forms of market structure – perfectly competitive and monopolistic forms – produced incentives that were less than socially desirable. However, Arrow's (1962b) analysis did not include a possibility of entry into the market as he considered this situation as "...more nearly competitive" (p. 619). Nonetheless, the mere threat of an entry can incentivize the monopolist to conduct *preemptive R&D*, and this was formally demonstrated in the game theoretic model of Gilbert and Newbery (1982).

The key to understanding of this phenomenon is found in patenting, yet the lack of patents does not entirely remove the incentive of preemptive R&D. A patent is essentially a right to exclude others from producing patented invention. In a continuum of innovations, this creates an incentive for the monopolist to conduct R&D in order to discourage potential entrants from developing substitutes to products of the monopolist. This can lead to a situations where patented inventions are neither produced or licensed (Gilbert & Newbery, 1982). However, the benefit from preemptive patenting is hampered by complexities of technological development, such as multiple competitive threats – variation in the number of possible research directions - and uncertainty in the invention process (Gilbert & Newbery, 1982). Thus, preemptive patenting is limited in its ability to prevent entry especially in industries with fast technological progress (Gilbert & Newbery, 1982).

Gilbert and Newbery (1982) also discussed implications of uncertainty for the behavior of agents in their model. According to them, preemption was incumbent's best response only when the *expected* gain from entry was positive for *every* challenger. Hence, the incumbent's preemption incentive simplifies to preempting the entry of the one that it expects to be the most optimistic entrant. Thus, expectations play the key role in preemption and therefore it can be concluded that institutions such as the patent system create opportunities for established firms with monopoly power to maintain their position, and hence, incumbent firms might have an absolutely stronger incentive to innovate certain kinds of innovations. (Gilbert & Newbery, 1982).

Next, for the sake of brevity, a reduced form of Gilbert's and Newbery's (1982) model à la Reinganum (1989) will be presented: If P^m denotes the value of being a monopolist, and P^e , P^i represent the entrants and the incumbent's proportions of the value of innovation, the replacement effect will be:

$$P^m \geq P^e + P^i$$

If the previous is true, then the incumbent will win the bidding game with a bid of x*, where x* is the largest solution of:

$$P^e e^{-rT(x)} - x = 0$$

, where r represents a discount factor, T(x) the arrival rate of invention, and x denotes the research intensity. The incumbent will preempt if:

$$P^m e^{-rT(x^*)} - x^* \geq P^i e^{-rT(x^*)}$$

, which simplifies to:

$$[P^m - P^i - P^e]e^{-rT(x^*)} \geq 0$$

This indicates that the incentive preemptive patenting arises from the dissipation of industry profits that are associated with increased competition. (Reinganum, 1989).

Reinganum (1983b) reinstated Arrow's (1962) conclusion about the innovative inefficiency of monopoly by adding innovation magnitude and stochastic innovation arrival rate to a similar analysis to that of Gilbert and Newbery (1982). The reason for these results was the same as in Arrow (1962b) – the *replacement effect*. Reinganum (1983b) shows that when:

> "*...the first successful innovator captures a sufficiently high share of the post-innovation market (i.e., when the innovation is sufficiently revolutionary), then in a Nash equilibrium the incumbent firm invests less on a given project than the challenger*" (p. 741).

What is more, Reinganum (1983b) shows that under an alternative specification of the model the incumbent firm will conduct fewer parallel projects than the challenger. According to Reinganum (1989) the logic for these results is a straightforward one:

> *"...at least for the case in which the innovation is drastic; that is, when the innovator captures the entire post-innovation market. When innovation is uncertain, the incumbent firm receives flow profits before successful innovation. This period is of random length, but is stochastically shorter the more the incumbent (or the challenger) invests. The incumbent has relatively less incentive than the challenger to shorten the period of its*

incumbency" (p. 871).

Reinganum (1983b) extends the former auction game of Gilbert & Newbery (1982) to a stochastic racing model. Let $\bar{c}$ be the incumbent's current unit costs and c the unit costs associated with a new technology, $c<\bar{c}$. Let $\Pi(c)$ be the present value of monopoly profits under new technology and R the current flow rate of profit. Finally, let $\pi_I(c)$ and $\pi_C(c)$ be the Nash-Cournot profits to the incumbent and the challenger, respectively, if the challenger innovates and the incumbent continues to use the old technology. Functions $\Pi(c), \pi_I(c),$ and $\pi_C(c)$ are assumed to be are continuous, and piecewise continuously differentiable. In addition, $\Pi(c)$ and $\pi_C(c)$ are assumed to be non-increasing in c, while $\pi_I(c)$ is non-decreasing in c. Again, innovation will be considered drastic if $c \le c^0$ where c^0 is assumed to exist and it is the largest value of c such that $\pi_I(c)=0$. If the unit cost is larger than c^0 , the incumbent will leave the market. The critical assumption is that there are constant returns to scale, which ensures that in the case of a drastic innovation incumbent's profits are zero, and so is its output. Another assumption, $c<\bar{c}$, ensures $\Pi(c)>R/r$, (r denotes the discount rate) and that $R/r=\Pi(\bar{c})>\pi_I(\bar{c})\ge\pi_I(c)$, which implies that innovation is always profitable in terms of present value. (Reinganum, 1983b).

Let a_i i = I, C denote the rival hazard rate $(h(x_i))$ for firm i, where x_i represents an individual R&D investment of either the incumbent (denoted I) or the challenger (denoted C). The hazard function (e.g. $a_I=h(x_C)$) is assumed to be twice continuously differentiable. The expected profit to the incumbent from being the first to innovate is:

$$V^I(x_I,a_I)=\int_0^\infty e^{-rt}e^{-h(x_I+a_I)t}[h(x_I)\prod(c)+a_I\pi_I(c)+R-x_I]dt$$
$$=\frac{h(x_I)\prod(c)+a_I\pi_I(c)+R-x_I}{r+h(x_I)+a_I}$$

, and for the challenger. (Reinganum, 1989):

$$V^C(x_C,a_C)=\int_0^\infty e^{-rt}e^{-h(x_C+a_C)t}[h(x_C)\pi_C-x_C]dt$$
$$=\frac{h(x_C)\pi_C(c)-x_C}{r+h(x_C)+a_C}$$

The main reason behind why agents have different payoffs is that the incumbent continues to receive a flow of profits R until the innovation is achieved, and that the incumbent receives a profit of $\pi_I(c)$ if the challenger is successful. A *strategy* for firm i is an investment rate x_i and a best response function for firm i is a function $\hat{x}(a)$ such that for all a, $V^i(\hat{x}_i(a), a) \geq V^i(x_i, a)$ for all x_i . A *Nash equilibrium* is a pair (x_I^*, x_C^*) such that $x_I^* = \hat{x}_I(a_I^*)$, where $a_I^* = h(x_C^*)$, and $x_C^* = \hat{x}_C(a_C^*)$, where $a_C^* = h(x_I^*)$. Thus, there exists Nash equilibrium pair $(x_I^*(c, R), x_C^*(c, R))$ [14]. (Reinganum, 1983b; Reinganum, 1989).

Moreover, according to Reinganum (1989):

> "*The incumbent's best response function is upward-sloping; thus the existence of the challenger provokes the incumbent to invest more than it otherwise would. If the innovation is drastic and R > 0, then in a Nash equilibrium the incumbent invests less than the challenger. That is,* $x_I^*(c, R) < x_C^*(c, R)$ " (p. 873).

A corollary to the previous analysis is if R > 0, it can be shown that there exists an open neighborhood next to the "division line" of drasticity, denoted $N(c^0; R)$, such that even if the innovation is not drastic, the incumbent will invest less than the challenge (Reinganum, 1983b). This follows directly from the continuity assumptions in equilibrium R&D investment rates $x_I^*(c, R)$, $x_C^*(c, R)$, and in innovation size c (Reinganum, 1983b). Figure 2.5 depicts this situation in terms of increasing innovation magnitude, e.g. cost reduction capability (illustrated by the arrow in figure 2.5). $\bar{c}$ denoted incumbent's pre-innovation unit costs, however, in this figure c indicates a drastic innovation, and c_0 indicates the lower bound of $N(c^0; R)$. Figure 2.5 illustrates Reinganum's (1983b) conclusions: The incumbent and the challenger will have differing incentives to invest in different innovations sizes. This makes it *unlikely* that each of the agents will win the race for a patent in either side of the division line of c_0.

According to the analysis presented above, innovation magnitude is the outcome of economic incentives which are in turn derived from prevailing market structure. Reinganum (1983b) was the first to suggest that her model was more suitable for the context of radical innovations while it would be more appropriate to associate incremental innovations with the model of Gilbert and

---

14 $x_I^*(c, R)$ is continuous in (c, R) for i = C,I

Newbery (1982). That is, Reinganum's (1983b) model had a stochastic inventive process while Gilbert's and Newbery's (1982) had a deterministic one. These assumptions change the expected payoffs associated with innovations. Furthermore, Reinganum's (1983b) model was a simultaneous move game while in Gilbert's and Newbery's (1982) game the incumbent had a first-mover advantage. This difference makes preemption possible (Gilbert & Newbery, 1984). Hence, what is essentially meant by the claim that radical innovations would be more suitable for Reinganum's (1983b) model is that these innovations involve uncertainty and a simultaneous order of moves. Conversely, incremental innovations involve *less* uncertainty (Gilbert & Newbery discussed this possibility) and the incumbent has a first-mover advantage in these innovations.

Figure 2.5: Illustration of R&D spending preferences in Reinganum's (1983b) model

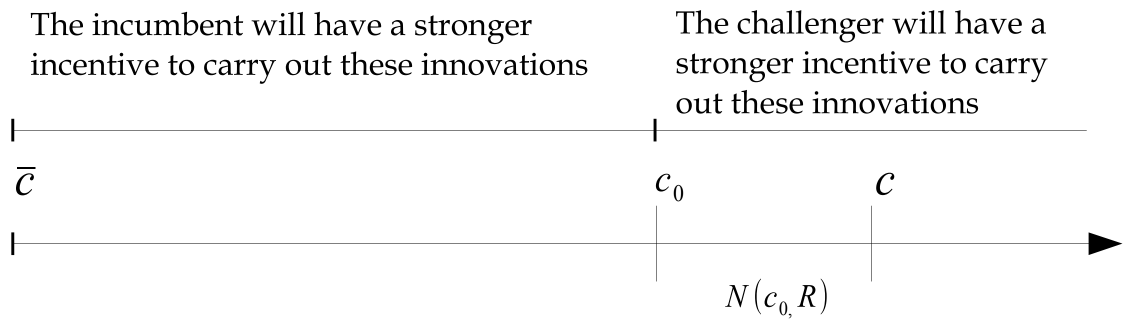


Table 2.3 summarizes the analysis above. The vertical alternative represents Reinganum's (1983b) result, which states that the incumbent will have a lower incentive to invest in R&D that is sufficiently large *if* this investment will hasten the innovation's arrival date. The horizontal axis, on the other hand, relays whether innovation is drastic or not.

Table 2.3: Radicalness measured by two criteria

	Implications for existing market power?	
Dependence on incumbent investment?	$c \le c^0$ [15] (Innovation makes the existing technology obsolete)	$c > c^0$ (Innovation competes directly with the existing technology)
Introduction date independent of incumbent investment	*?* Incumbents and entrants have equivalent incentives to invest	*Incremental innovation* Incumbents have a greater strategic incentive to invest
Introduction date a function of incumbent investment	*Radical innovation* Entrants have the greater strategic incentive to invest	*?* Investment behavior depends on the balance of incentives

Adopted from Henderson (1993. p. 250)

15 In $c \le c^0$ c^0 is the largest value of c such that $\pi_I(c)=0$, i.e. the incumbent will exit the market.

2.4 Linking technological evolution with industrial dynamics – a definition of an incremental innovation

First and foremost, it is necessary to distinguish invention from innovation: Arthur (2007) defined an invention as a process that either connects a need with an effect or which starts with a discovery of a novel effect to which a need suggests itself. This is what births a new technology – one which can be thought of a theoretical unit which gives s structure to a cumulative path of elementary small innovations (Zuscovitch, 1986). Such generic technology is depicted in figure 2.2 and it includes a need, a device, and a body of knowledge that is required to connect the two. Hence, the generic technology can be studied from three paradigmatic starting points.

Firstly, a technology can be seen as *a body knowledge*. In this view, technological knowledge can be identified with scientific knowledge (Dosi, 1982). A technological paradigm defines an artifact's base principle[16], its purpose, and the perceived opportunities of improving the device which originate in positive and negative heuristics denoted by the underlying paradigm. According to this view, an incremental innovation one that uses the same base principle as a previous radical invention, and thus, an invention which follow the establishment of a technological paradigm (a radical invention).

Secondly, in addition to the previous, a technology can be seen as the material *device* itself. In this view, technology can be thought of as a complex system (Murmann & Frenken, 2006). According to this idea, an incremental innovation is one that concerns the peripheral components and one that doesn't change the underlying system design (the arrangement of core components) (Henderson & Clark, 1990; Murmann & Frenken, 2006).

Thirdly, technology can be seen as filling a *need* of some sorts (Zuscovitch, 1986; Arthur, 2007). This need rise from economic opportunity, a social change, or it can be a military one[17] (Arthur, 2007). All that this paradigmatic starting

---

16 Operational principle can be defined for example as an method of cutting grass: a scythe vs. a lawnmover vs. an electric lawnmover – all fill the need for a smoothed sward.

17 The writer feels obliged to direct the reader's attention to the surrounding world and the fact that innumerous truly novel technologies have either partly or entirely originated in militaristic needs, e.g. cell phone, integrated circuit jet engine, nuclear power, rocketry, tea bags, wristwatches zippers, stainless steel, penicilin and so forth.

point enables to say about incremental innovations is that these innovations answer to an existing need – one that has been recognized by a previous (radical) invention.

On the other hand, game theoretic industrial organization literature focuses its examination to particular innovation related situations. Symmetric models offer rather generic results, e.g. that the amount invested by an individual firm decreases with the number of firms yet the aggregate industry R&D spending increases with the number of firms, and in the resulting equilibrium, with unrestricted entry, there will be excess R&D capacity (Reinganum, 1984b). However, existing asymmetric models can be used to examine incremental innovations. Gilbert and Newbery (1982) modeled monopolist's behavior in a game situation with first-mover advantage and a deterministic innovation process. With this initial set up, one who wins the race for a patent (i.e. overspends the rival) is the one who has the most to gain from doing so. Reinganum (1983b), on the other hand, modeled monopolist's behavior with a simultaneous move game with stochastic innovation process. With this this initial set up, the innovation's arrival rate is of random length but stochastically shorter the more the incumbent (or challenger) spends on R&D (Reinganum, 1989).

The analysis of technology centered innovation and technology literature (chapter 2.2) suggests a compatibility with firm centered industrial organization literature (chapter 2.3). The link can be found in varying levels of uncertainty before and after the emergement of a dominant design, i.e. in eras of ferment and of incremental change (Anderson & Tushman, 1990). The era of ferment involves larger uncertainty about possible research and design directions. This is because the design of core *and* peripheral components are in lucid motion of thought. The emergement of a dominant design solidifies the arrangement of core component and forces the problem solving activities to shift to lower pleiotropy components (Murmann & frenken, 2006). This changes the form of technological uncertainty faced by competing firms since future technological developments aim to extend and improve the dominant design. Therefore, is feasible to associate stochastic innovation process with the era of ferment and deterministic innovation process with the era of incremental change (Reinganum, 1983b). Murmann & Frenken (2006) proposed a similar dynamism:

"Before a dominant design emerges, for example, many small firms engage

in product innovation in an attempt to develop and appropriate the knowledge underlying the future dominant design, or parts of it. After a dominant design emerges, and the industry has become concentrated, firms have an incentive to engage in process innovations, the costs of which can be spread over many products (Klepper, 1997). This means that changes in technology and market structure are causally connected." (p. 945).

The combination of evolutionary dynamics with industrial outcomes is indicative of several derivative conclusions. Firstly, one could expect that incumbency changes hands far more often in the era of ferment in accordance with assumptions and the model of Reinganum (1983b). Moreover, one could also expect the number of firms to be larger than in the era of incremental change, and hence, the aggregate R&D spending, as well. This claim is based on symmetric innovation models (more in Reinganum, 1984b; Reinganum, 1989). The emergement of a dominant design changes the uncertainty to more deterministic kind, and thus, makes preemption possible. Secondly, one should expect monopolistic competition from the era of incremental change in accordance with the model of Gilbert and Newbery (1982). Moreover, one could also expect lower aggregate R&D spending in this era in accordance with symmetric models (more in Reinganum, 1984b; Reinganum, 1989).

From this theoretical background it is possible to synthesise a definition for incremental innovation. While radical innovations can take, perhaps, an infinite number of forms in the modern world as they are novel *ideas* (see Weitzman, 1998), incremental innovations are destined to refine and and extend an established design that utilizes a given base principle. Or put differently, *incremental innovation is the refinement of the underlying device*[18] in Zuscovitch's (1986) triangulum, while radical inventions establish the triangulum. Hence, incremental innovations are ordered and confined, and this can be thought of as a technological trajectory (Dosi, 1982).

If, on the other hand, technologies are thought of as complex and hierarchical systems, incremental inventions would simply relate to the lower levels of a hierarchy (Murmann & Frenken, 2006). However, there is one problem with this definition: the measure of incrementality depends on the investigated entity. An improvement to a subsystem is an incremental innovation from the perspective of the whole system, while it can also be a

18 As incremental innovations do not require external knowledge, such as scientific inputs, they could be characterized as "iterated" problem solving activity (Dosi, 1982). A process that might also be called *learning*. More in appendix B.

radical invention from the perspective of some subsystem. Hence, it must be concluded that innovation's incrementality increases when it takes place closer to the component level. That is, innovations become *more* incremental the lower the hierarchical system level that they belong to. This in turn is likely to happen after the emergement of a dominant design[19] - i.e. the solidification of the arrangement of core components to a singular design.

Nevertheless, as is evident from the description until now, incremental innovations can only take place in certain kinds of industries and certain kinds of products. Firstly, the product in question needs to be an *artifact* of some sort, i.e. the product needs to be *tangible* and there needs to be room for the refinement of this artifact. This requirement relates to the uniformity of the base principle (what to do and how exactly to do it). If for example a drug is replaced with a more effective one, this essentially amounts a new effect and, hence, a new base principle[20]. Secondly, the formation of a dominant design is by now means a predetermined phase in every technology's evolution (Dosi & Nelson, 2013). It is entirely possible that a technology is supplanted by another before a dominant design can emerge. Furthermore, the causal logic leading to the emergence of a dominant design is not well understood (Murmann & Frenken, 2006). Therefore it must be concluded that incremental innovation as a concept and innovation typology as a whole is at best Kuhn's (1962) conceptual paradigm in the development of science.

19 Technological paradigm can either be thought of as the generic technology of Zuscovitch (1986) or as the structure that a dominant design brings to bare in the problem solving activity (Murmann & Frenken, 2006). If these two views are synthesized, one could simply state that the technological paradigm drastically changes its form when a dominant design emerges

20 According to this example and the definition of incremental innovations that is presented here all innovations that take place in the pharmaceutical and chemical industries are to a some decree radical.

3 CRITICAL REVIEW OF THE PRIOR EMPIRICAL LITERATURE

3.1 Methodological issues related to the empirical study of innovation typology

The last section synthesized existing technological and industrial organization literature in order to construct a coherent picture of what technology actually is, how it evolves, and what incremental innovations might be in the grand scheme of technological evolution. This section reviews the existing empirical literature in order to establish methodological guidelines for the study of technological change – from an incremental standpoint. Various scholars have studied innovation typology empirically prior to writing this thesis. The focus of these, and approximately any innovation typology related study, seems to be the identification of radical inventions and innovations. This is not too surprising considering that these innovations – unlike incremental ones – are destined to result in large changes in the competitive landscape of an industry. Hence, this section must approach empirical innovation typology literature from a radical standpoint.

According to Dahlin and Behrens (2005) various scholars have used different measures of radicalness. Thus, it is necessary to identify general guidelines for assessing innovations' radicalness or incrementality. Here it will be proposed that there are actually three dimensions to an invention's radicalness. Firstly, Verhoeven, Bakker and Veugelers (2016) point out the fundamental need to distinguish between ex ante technological characteristics of invention and its ex post *technological*[21] impact, as these two measuring criteria do not necessarily highlight the same inventions as equally radical. In addition, assessment of an invention's radicalness should distinguish between the ex post impact of an invention and the ex post impact of an innovation. In order to reiterate previous statements, there are three things to consider when assessing radicalness: invention's antecedents (i.e. the novelty of the base principle and invention's hierarchical level), its technological impact (i.e. the way a particular invention affects future inventions), and a subsequent innovation's market impact (i.e. how the invention changes the competitive landscape within an industry). This taxonomy is illustrated by Figure 3.1. Any method of measuring an invention's radicalness – or incrementality – must be contrasted with these three criteria. At the same time, it should be remembered that these three criteria of innovations' radicalness are essentially part of a unity – i.e. the (tangible) innovation, it and how it affects the world.

In an influential and recent paper, Dahlin and Behrens (2005) proposed the following three point criteria for measuring inventions' radicalness:

"*Criterion 1: The invention must be novel: it needs to be dissimilar from prior inventions.*

Criterion 2: The invention must be unique: it needs to be dissimilar from current inventions.

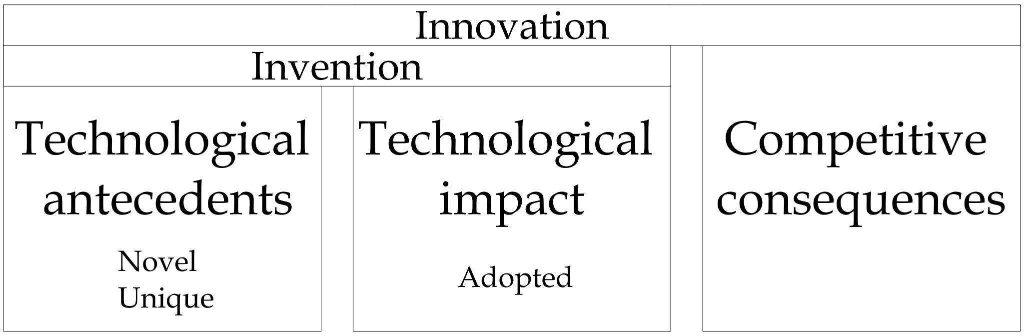
Criterion 3: The invention must be adopted: it needs to influence the content of future inventions" (p. 725).

Dahlin and Behrens (2005) propose that if an invention fills the first two requirements but not the third one, can be described as a *failed radical invention*: an invention that is intrinsically different from past and current inventions, but fails to change the course of future inventions. That is, an invention's radicalness depends on three moments in time: past, present and future (Dahlin & Behrens, 2005). In the light of the discussion above, criterion 1 and 2 are parts of inventions' antecedents and criterion 3 can be associated with invention's technological impact. Thus, novelty and uniqueness can be included in the concept of invention's technological antecedents, while adoption relates to its technological impact. Figure 3.1 illustrates.

---

21 This conept is not to be mistaken with the concept of drasticity. The ex post technological impact of invention relates to the changes in the onging search process for new inventions, i.e. technological trajectories (Dosi, 1982). while drastic *innovations* relate to changes in the market stucture.

Figure 3.1: Taxonomy of the dimensions of radicalness



In a recent article, Verhoeven et. al. (2016) use a different kind of approach to technology and radicalness. In their view, technologies differ from each other in recombination of existing components and principles, and hence, inventions' radicalness can be measured with the *novelty* of such recombinations. Moreover, technological novelty, as proposed by Verhoeven et. al. (2016), can be fragmented into three kinds of novelty: novelty in recombination, novelty in technological knowledge origins, and novelty in scientific knowledge origins. According to them, novelty in recombination means that the combination of components and base principles is new to the world, novelty in technological knowledge origins means that an invention draws knowledge from domains that were not previously drawn upon by the domain of the invention, and novelty in scientific knowledge origins means that an invention draws on scientific domains that were not previously drawn upon by the domain of the invention (Verhoeven et. al., 2016). However, it is clear that this method also belongs to inventions' technological antecedents.

Now that the reader is familiar with general methodological issues related to measuring inventions' radicalness, it is possible to review the individual methods. The following chapter will review non-patent based and patent based methods of measuring radicalness. The subsequent chapter will discuss the usability of these methods in the measurement of incremental innovations.

3.2 Measuring radicalness

There are altogether three generic non-patent based methods of measuring inventions' radicalness: performance increases, hedonic price models, and expert panels (Dahlin & Behren, 2005). First, the radicalness of innovations can be measured with the *performance increases* that they introduce, e.g. with the acceleration of a new car model, or with the amount of paper produced per hour by a new paper-making machine. Tushman and Anderson (1986) and Anderson and Tushman (1990) used this method to distinguish technological discontinuities: When an innovation can be connected with a large observed performance increase, such an innovation can be termed to be radical. Dahlin and Behrens (2005) criticize this method for its reliance on unidimensional performance improvements while it is entirely possible that a radical innovation creates a new performance criteria, as Dosi's (1982) ideas of technological paradigms suggest. On the other hand, Dahlin and Behrens (2005) criticize Dosi (1982) for presenting a view where radical innovations can only be observed ex post. In addition, a novel technology is likely to underperform the old technology for a period of time after its invention (Dahlin and Behrens, 2005; Rosenberg, 1976). Overall, these problems will lead to under-reporting of radical innovations (Dahlin and Behrens, 2005).

Second, the radicalness of innovations can be measured with *hedonic price models* (e.g. Henderson, 1993). These regression models use product price as a dependent variable and product characteristics as independent variables. If an innovation is radical (or drastic in this case), one could expect it to command an unique price premium (Henderson, 1993). However, according to Dahlin and Behrens (2005), this method can be criticized for its sensitivity to model specification: Firstly, a researcher must understand the technology in question so that the product characteristics can be specified. Secondly, regressions must include various product characteristics in order to produce a robust and consistent result of radicalness (Dahlin and Behrens, 2005). Thirdly, it is dubious that market's willingness to pay more can be traced directly to radical innovations as it is entirely possible that small incremental improvements might also lead to large price increases (Dahlin and Behrens, 2005). Overall, hedonic price models essentially connect the drasticity of an innovation, i.e. its competitive consequences in the marketplace, to its technological antecedents. Yet, this approach requires encompassing knowledge about the industry and the technology in question, and is hence a cumbersome tool in the empirical study of innovation typology. In addition, it should be noted that hedonic price models are useless as forecasting tools (Dahlin and Behrens, 2005).

Third, *expert panels* (e.g. Pavitt, 1984). Individuals concerned with studying a particular technology surely know what developments can be called radical and which ones incremental. Expert panels, on the other hand, suffer from human gullibility and biases that result from it, i.e. success and availability bias (Dahlin & Behrens, 2005). Success bias means that experts might be more acknowledging to innovations that have done well in the marketplace and they might also rate these innovations more favorably (Dahlin & Behrens, 2005). Availability bias, on the other hand, means that experts are likely to emphasize information that is closer to their past and present experience (Dahlin & Behrens, 2005). Hence, it must be concluded that expert panels are a laborious and possibly a biased method of assessing radicalness or incrementality.

Based on the three generic non-patent based methods discussed above, it can be argued that the most applicable methods of conducting large scale quantitative assessments of inventions' radicalness involve the use of patent statistics. The advantage of patent data is that it does not suffer from retrospective or success biases, that plague expert panels, as it is produced continuously and prior to the commercialization of the underlying invention (Dahlin & Behrens, 2005). Patent data also enables a large scale assessment of inventions as it is easily accessible and computational (Verhoeven et. al., 2016). On the other hand, the consistency of intra-industrial patenting rate depends on the size and diversification of the whole economy (Cohen, Nelson & Walsh, 2000; Nikulainen, 2008). Thus patent data needs to be contrasted with its industrial origin.

A patent application contains four kinds of information: technical details of the invention, information about the applicant, administrative identifiers (such as international patent classifications, i.e. IPCs[22]), and citations to other patents. Technical details enable the recreation of the invention, information about the applicant enables associations with industries, identifiers such as the IPC aid in the administrative procedures associated with the registration process, while patent and science citations are used to limit the scope of the exclusion right. Two of these information kinds can be used measure radicalness: citation information (e.g. Dahlin & Behrens, 2005) and administrative information contained in IPCs (e.g. Verhoeven et. al., 2016). Hence, there are also three generic patent based methods of measuring

22 IP classification system is discussed in greater detail in chapter 4.2.

inventions' radicalness: backward citations, forward citations, and patent classification analyses.

Citation analysis can be directed either backward and forward. Backward citation information refers to cited material in a patent application while forward citations are future references to this particular application. The existence of a stream of citations can impart information about a continuing development of the underlying technology, and hence, hint at the existence of a market for the invention (van Zeebroeck, 2011). Furthermore, if a patent is used by a patent examiner to limit the scope of a future patent application, it can be deduced that the cited patent has social value (van Zeebroeck, 2011).

Dahlin and Behrens (2005) criticize patent citation related measurements of radicalness for not being pure measurements of technological content. Stuart and Podolny (1996. Cited in Dahlin & Behrens, 2005) found that that a patent is more likely to be cited in future if the patent owner has patented many other inventions in the past. Hence, a citation can also be a measurement of applicant's social status (Dahlin & Behrens, 2005). However, these problems relate to *forward citations*. Backward citations, on the other hand, can be a useful tool for measuring radicalness. The idea is that patents that cite scientific sources or unusual patent classes are likely to be more radical (Dahlin & Behrens, 2005). The latter backward citation related method concerns the informative content of IPCs: A radical inventions might be more likely to cite patents from other patent classes than to which the invention belongs to (Rosenkopf and Nerkar, 2001). Moreover, Shane (2001) simply counted the number of patent classes that were cited. The idea was the following: the more classes were cited, the more radical an invention was.

The most comprehensive method of measuring radicalness using backward citations was developed by Dahlin and Behrens (2005). As was already mentioned, they divided inventions' radicalness into three sections: novelty, uniqueness, and adoption. Novelty was measured by the uniqueness of backward citation patterns, uniqueness was also measured with backwards citation patterns but the comparison was found in patents filed in the same year as the focal patent, and lastly, adoption was measured by a comparison of the focal patent citation structure with future patent citation structures. However, Verhoeven et. al. (2016) criticize the method developed by Dahlin and Behrens (2005) of its computational complexity and of difficulties associated with selecting a comparison group in case of analysis with multiple multiple technological fields. Dahlin and Behrens (2005) were able to use this method on a small sample of a specific technology – the tennis racket. Yet, this method is still to be used on large scale and cross technologies (Verhoeven et. al., 2016).

Nevertheless, it is also possible to simply use the information contained in IPCs to make assessments of inventions' radicalness without a formal citation analysis. Lerner (1994) used the *number* of IPC subclass (four digit) symbols as a proxy for patent scope, i.e. the importance of a patent: The more IPCs a patent included, the larger the scope it had. Gruber, Harhoff and Hoisl (2013) used the *combinations* of different IPCs as proxies of technological novelty: According to them, inventions that span technological boundaries in novel ways are more likely to be radical. However, the establishment of "technological boundaries" requires a statistically useful taxonomy of IPCs and industries (Gruber, Harhoff & Hoisl, 2013). A similar, yet more refined, method was presented by Verhoeven et. al. (2016).

As was mentioned before, Verhoeven et. al. (2016) divide patent's novelty into three factors: novelty in recombination, novelty in technological origins, and novelty in scientific origins. Novelty in recombination means that a patent contains at least one pair of IPC groups that have not been combined before in a patent application. Novelty in scientific origins means that a patent makes a connection that with a scientific source that has not been done prior to the application year. Novelty in technological origins, on the other hand, is measured by a comparison of backward citation pairs, i.e. IPC(s) of the patent and IPCs of cited patents: The patent has novelty in technological origins if its IPCs are not similar to the patents that it cites.

To summarize, methods of measuring invention's radicalness can be divided into two broad categories: non-patent based and patent based ones. Non-patent based methods include performance increases, hedonic price models, and expert panels. Patent based methods include backward citations, forward citations, and patent classification analyses. Table 3.1 summarizes the discussion above and adds an additional question about methods' usability in assessing whether an method is applicable in the study of incremental inventions and innovations. It asks on which dimension of radicalness a method measures, does a method involve selection bias, how available is data, and whether a method is suitable for large scale analysis. In addition, it asks whether a method is suitable for the measurement of incrementality. This last question will be further elaborated in the following chapter.

Table 3.1: List and comparison of methods of studying innovation size

Summary of generic methods of measuring radicalness that have been used in empirical innovation typology literature

	Performance increases	Hedonic price models	Expert panels	Backward citations	Forward citations	Patent classification analyses	Combination of patent statistics
Granularity of analysis[23]	B) C) Technological consequences and/or market consequences	C) Market consequences	A) B) C) Possibility of assessing innovation as a unity; every level of analysis possible	A) Analysis is backward-looking	B) Analysis is forward-looking but doesn't *necessarily* tell about market consequences	A) Analysis only deals with the focal patent application	A) B) Combination of methods removes many patent statistics related short-commings
Selection bias	Yes – only highlights successful innovations	Yes – uses only characteristics of successful innovations	Yes – results vulnerable to human gullibility	No – tells whether an invention is *novel*	Yes – patent applicant's social status might matter for citation	No – tells whether an invention is *novel*	No – tells whether an invention is novel and adopted
Availability of data	Linking market outcomes to innovations can be arduous	Experts needed to assess technological antecedents	Can be expensive	Data easily available	Data easily available	Data easily available	Data easily available
Possibilities for large scale quantitative analysis	No – data acquisition can be arduous	No – data acquisition can be arduous	No – data acquisition can be arduous, difficult, and expensve	Yes	Yes	Yes	Yes
Suitability for the measurement of incrementality	Unsatisfactory – no clear divergence between success. unsucess innovs.	Unsatisfactory – it is possible that an inc. Innovation results a large price difference	Avarage – expert can assess incrementality but are vulnerable to e.g.success bias	Unnecessary – incremetal inventions can be novel	Unnecessary – inc. inventions' technological consequences are irrelevant	Unnecessary – incremetal inventions can be novel	Unsatisfactory

23 This denotes whether the method applies to A) invention's antecedents, B) technological consequences, or C) innovation's market consequences.

3.3 Measuring incrementality

As is evident from the table 3.1, none of the methods that were reviewed in the last chapter are especially suitable for the study of incremental inventions and innovations. This is not too surprising considering that these methods were developed for the study of radical inventions. To my knowledge, no other paper have been written about incremental innovations, in particular, before this one. Hence, there is no existing method that is especially suitable for separating incremental innovations from a large group of inventions. This is not too problematic if the focus of a study is in identifying radical inventions: the issue becomes dichotomous and inventions that are not radical, can be called incremental.

Yet, the idea of a technological evolution or a technological continuum that was depicted in section two advises not to view innovation typology is such simplistic terms. Furthermore, the radicalness of an innovation can be separated into three dimensions: technological antecedents, technological impact, and innovation's market impact. As it happens, existing methods of assessing innovations' radicalness assess some of these dimensions, but not all of them[24]. That is, existing methods can be criticized of being unidirectional and inadequate assessing the unity of radical innovations (antecedents, technological impact, and market impact).

Performance increases and hedonic price models reveal innovation's incrementality thorough a proxy (performance or price), and hence, they do not assess the related typology directly. Expert panels, on the other hand, is a cumbersome tool for large scale statistical analyses. Conversely, patent statics enable large scale computational analyses with relative ease. However, studies that employ patent statistics (e.g. Dahlin & Behrens, 2005; Verhoeven et. al., 2016) focus on radical inventions. A method of assessing invention's incrementality could use the contrary conclusions of these existing methods, yet it is also possible to create a method that is especially suitable for the study of small innovations. This is the topic of the next section which will present one more method that can be used to study incremental inventions – the utility model. This section will function as a guideline for the development such a method.

24 That is, with the exception of expert panels. This method, however, can be biased (availability and sucess bias), expensive, and laborous method of assessing innovations.

4 EMPIRICAL ANALYSIS

4.1 The utility model

Although incremental innovations can be studied with the aforementioned methods, there is also another method which is explicitly suitable for the measurement of incremental innovations: *the utility model*. The idea is that this "petty patent" can function as a proxy for an incremental invention[25]. Beneito (2006) applied this method to a research input-output analysis, where she considered patents and utility models as approximations of radical[26] and incremental innovations. To my knowledge, this is the only occasion where this method has been employed prior to writing this thesis.

Utility models differ from regular patents in several aspects: First and foremost, the utility model has a less stringent patentability requirement than regular patent which means that it has a larger probability of being granted. In addition, it should be noted that the utility model also implies a shorter statutory patent life, absence of yearly renewal fees, lesser overall application costs, and significantly faster and simpler registration procedures (WIPO). However, utility models can in some instances be revoked without a formal court decision, making it potentially less safe than a regular patent. All in all,

25 Incrementality of utility models is examined in appendix C with a case study of Finnish paper-making machine industry.

26 The term that was used in this study was *significant* which is not exactly the same as a radical invention. What is notewhorthy here is that the taxonomy that was used in Beneito (2006) was *dichotomous*.

according to World Intellectual Property Organization (WIPO), utility models are sought especially for the protection of inventions “of a rather incremental character which may not meet the patentability criteria", and they are thus “particularly suited for SMEs that make minor improvements to, and adaptations of, existing products” (WIPO).

However, this study focuses on Finnish utility models. The not so stringent patentability requirement of Finnish utility models is evident in the wording that is used to describe these requirements: the requirement for inventive step is replaced with distinctiveness while the other requirements of novelty and industrial applicability remain the same (PRH, 2012). In addition, it must be stated that the Finnish patent and registration office does not inspect utility model applications unless they are explicitly challenged. Moreover, the statutory patent life of Finnish utility models is 10 years, and this needs to be contrasted with the statutory patent life of patents which is 20 years. Furthermore, Finnish utility models do not involve yearly renewal fees, and thus, the cost of a consummated utility model (10 years) with all costs included is less than 1000 euros. In contrast, registration costs of a Finnish patent are approximately as much, however, yearly renewal fees make a patent much more costly. And lastly, Finnish utility models can be received fast; the registration procedures take approximately three months (PRH, 2012).

What is more, it should also be stated that utility models are awarded only for product and never for process inventions. This means that in order to be awarded an utility model, an invention must constitute a device. This shifts the emphasis of utility models away from such industries as pharmaceuticals and chemicals. In addition, it should be noted that it is possible to patent an invention with a patent and an utility model, and it is possible to transform a patent application into an utility model application, but not vice versa.

Utility models as a method of identifying incremental inventions is a straightforward one as the power of using patent type as an indicator of innovation type resides in the usefulness of *patent counts*. The idea is the following: When there are two kinds of patents in existence, if an invention is perceived ex ante to have a low level of validity, its inventor will select the lower tier patent, irrespective of invention's economic value (Atal & Bar, 2014). Hence, lower tier patents, i.e. utility models, can be expected to be inventions of low perceived validity which means that they are perceived to be rather irrelevant inventions or small improvements to an existent whole, i.e. incremental inventions. Thus, it can be expected that utility models represent a constrained set of inventions: small and smaller, and therefore, it would be more appropriate to simply count the utility model applications.

Further arguments that support utility model's applicability to function as a proxy for an incremental invention are the following. Incremental inventions might be so 'small' that they do not fill the patentability requirements and are hence targets of the utility model (Beneito, 2006). Moreover, the absolute novelty requirement, which is ensured by a formal inspection conducted by patenting authorities, guarantees the inventiveness of the utility model. Furthermore, inventions that are protected by utility models are tangible inventions which means that utility models are in the same category as patents when compared with e.g. design patents: they are both *technologically* inventive.

On the other hand, utility models are cheaper than patents which can meant that if both forms of legal protection are assumed to be as stringent, financially constrained firms and individuals would choose the utility model regardless of the perceived invention size. In addition, as it is possible to protect an invention with both forms of legal protection, firms that do not face large financial constraints might choose both forms of protection because of the speed at which it is possible to receive a fully fledged utility model. Furthermore, it is entirely possible that patents might as well be incremental improvements to an existent whole. Nevertheless, it is clear that these three issues are not entirely sufficient to hamper the effectiveness of utility models as proxies for incremental inventions.

This section will continue as follows: The next chapter will construe key technology and industry classification systems (IPC and NACE) that are used in the following chapters. The following chapters will present descriptive statistics of technological content (IPCs) and industrial associations (NACEs) of Finnish utility models contrasted by the technological content and industrial association of Finnish patents. What technologies are more incremental? What industries are more incremental? The empirical study of utility models will then continue with regression analysis of industrial utility model usage. What industrial characteristics explain observed utility model counts? The aim of this empirical part is to understand what part the theoretical concepts that were developed in section 2 play in the contemporary Finnish economy.

4.2 Taxonomy of technological content and industrial divisions

All European patent and utility model applications are assigned International Patent Classifications (IPCs) in order to distinguish technologies from each other. This practice was first initiated by the Strasbourg Agreement in 1971 and continues to this date. World Intellectual Property Organization (WIPO) ensures that IP classifications are up to date by periodical screenings. Overall, when compared with United States UPSTO classifications, the IPC system can be considered relatively more *industry* and profession orientated (Lerner, 1994).

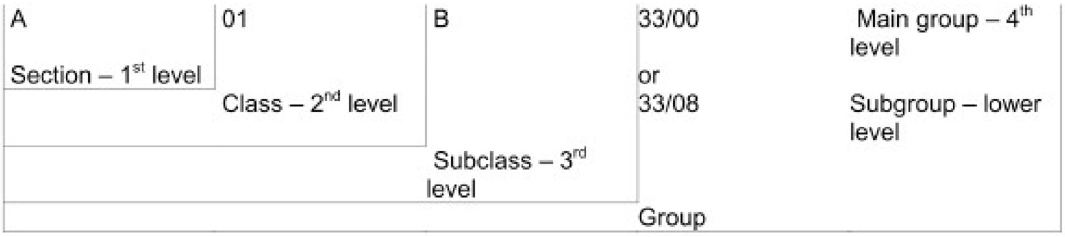
IPC is a *hierarchical* system which is made up of up to 7 digit long codes. A decomposition of an IP classification is illustrated in figure 4.1. The first level consists of a class symbol which is one of eight possible capital letters listed below.

A HUMAN NECESSITIES
B PERFORMING OPERATIONS; TRANSPORTING
C CHEMISTRY; METALLURGY
D TEXTILES; PAPER
E FIXED CONSTRUCTIONS
F MECHANICAL ENGINEERING; LIGHTING; HEATING; WEAPONS; BLASTING
G PHYSICS
H ELECTRICITY

The second level of the hierarchy (subsections) consists of 2-digit number which is followed by another capital letter, and together, these three hierarchical levels make up the *subclass symbol*. IPC subclass symbols denote the main technological fields, and hence, these symbols are only loosely connected with each other. However, each subclass symbol can be further divided into subdivisions called groups. Groups can be either main groups (denoted by a '/00' symbol) or subgroups, where each main group symbol can be particularized by various subgroup symbols. The existence and the number of subgroup symbols depends on the nature of the technology in question[27].

27 It is noteworthy that in some cases these hierarchies can be commensurate to the system hierarchy discussed in the definition of an incremental innovation (ch. 2.4), but the possibilities of parity are restricted to certain kinds of technologies which are are mainly complex mechanical devices.

Figure 4.1: Hierarchical nature of the IPC



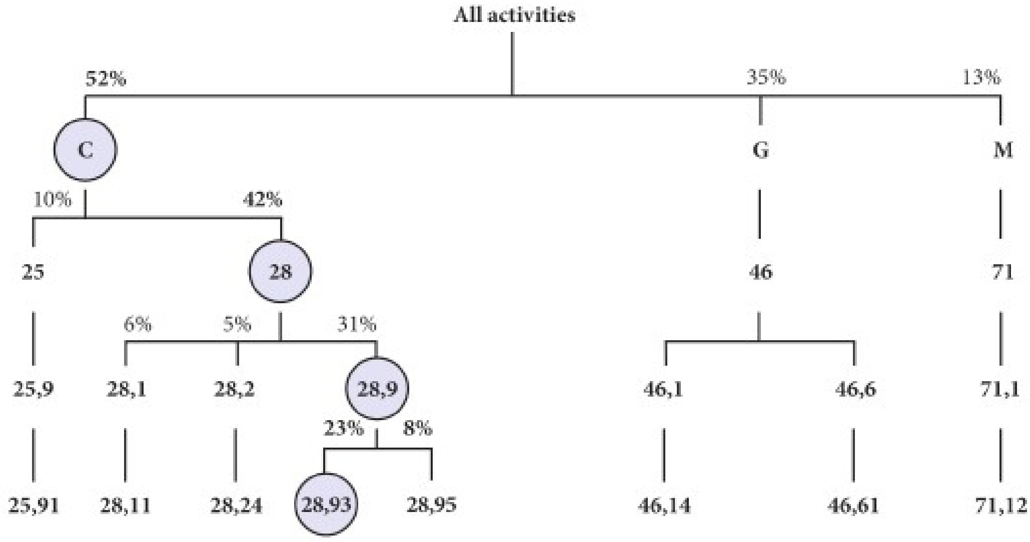
Adopted from IPC(version 2015, p. 6)

These classifications, however, tell nothing about the origin of an invention. Another classification scheme is needed for this purpose: Industrial origins of inventions can be segregated according to the NACE industry classification scheme[28], which is a hierarchical classification scheme of economic activities. NACE, short for Nomenclature générale des activités économiques dans les Communautés Européennes, was first introduced in 1970 but the scheme has since experienced multiple revisions and updates.

The most up to date version of NACE (Rev 2.) consists of four levels: The highest level (sections) is represented by an alphabetical code, and the lower levels are represented by a numerical code, which can be up to four digits long. These lower levels are: divisions (2 digits), groups (3 digits), and classes (4 digits). Figure 4.2 illustrates an exemplary NACE classification. The actual technology used in the production becomes less important on higher aggregated levels of NACE, which is to be expected that more general classifications of industries involve an increasing number of specific technologies. Moreover, divisions and groups are delineated according to the character of the goods produced, the uses to which the goods are put, and by the inputs, processes and technologies employed in the production of these goods (NACE Rev. 2, 2008). That is, NACE classifications classify *activities,* and hence, they relay information about the actual origin of an invention (individual firm or individual person), through a proxy of the aggregated industry.

28 NACE is the reference system for Finnish national version of NACE (TOL2008) where classifications are linked by a common structure. NACE is derived from ISIC (United Nations' International Standard Industrial Classification of all Economic Activities), but it is more detailed on lower levels.

Figure 4.2: Illustration of an exemplary hierarchical NACE classification[29]



Adopted from NACE Rev. 2 (2008. p. 29)

These two classification schemes – the IPC and the NACE – can be linked according to concordance schemes developed by Schmoch et al. (2003), which was later refined by Van Looy, Vereyen & Schmoch (2014). Hence, it is possible to link individual utility models to technologies and these technologies to industries. However, there is one problem associated with this: Utility models and patents usually include multiple IPCs. Available concordance tables (the one used in this study was acquired from PATSAT online) can only transpose singular IPC subclass symbols into NACE classifications. Hence, it is impossible to associate all patents and utility models with singular NACE industries with absolute certainty.

Yet, it is possible to study the technological content of both patents and utility models. This analysis should shed light on whether utility models can function as a proxy for incremental inventions. In addition, it is possible to measure the use of utility models and patents in varying industries. This in turn, should shed light on the "location" of incremental innovations within an economy, grated that utility models can function as a proxy for incremental innovations.

29 28.93 represents the manufacture of machinery for food, beverage and tobacco processing.

4.3 Data sources and variable definitions

This study utilizes patent, utility model, and industry data. First, patent data was acquired from the PATSTAT database. The search parameters looked for approved Finnish patents and associated IPCs from the period 2000-2013. The search produced 10 815 unique patent applications with 39 841 individual IPCs. Second, Finnish utility model data from period 1992-2014 was received directly from PRH and it included 16 766 unique applications with 25 831 individual IPCs. However, the removal of duplicate and rejected applications, as well as, restriction of sample interval to period 2000-2013 left 5762 unique utility model applications with 6 008 individual IPCs. Third, industry data was retrieved from Statistics Finland (STATFI) and it included Finnish industry specific firm count, turnover, personnel, and wage data from 2007.

Individual IPCs were assigned weights according the number of IPCs that were included in a particular patent or utility model. The total weight of each patent and utility model sums up to one, and hence, the number of IPCs denotes the weight fraction of each IPC and all IPC weights in a data set sum up to the number of total applications within a data set. Essentially this means that the weight of an individual IPC lessens the more IPCs a patent or an utility model contains.

Furthermore, few skims through the utility model data drew my attention to differences amongst inventions of firms and individuals. Hence, a decision was made to separate utility model applicants into firms and individuals. Utility model applications that contained one or several of the following were classified as potential Finnish companies, Oy, Ky, inc., corporation, limited, ltd. Individual applicants were then assorted with the incidence of comma in the applicants name (comma was only used to separate last names from first names and did not appear in any company's name). These measures successfully separated the 5762 utility model applications into 3 326 firm and 2 436 individual applications.

All patent and utility model IPCs were then transposed to a single list which included 16 986 separate IPC sub-classes, main groups, and subgroups, and these IPCs came with associated weight counts from patents and utility models. Correlations on table 4.2 are calculated from this data set. Next, all IPCs were abbreviated to IPC subclass level and combined to a single list. These measures shortened the IPC list to 575 subclass symbols. Correlations on table

4.3 are calculated from this data set.

These 575 subclass symbols were then converted to NACE.Rev.2 industry classifications by utilizing a IPC-NACE conversion table that was acquired from PATSTAT. This conversion table is based on the original work of Schmoch et al. (2003) and updated version of Van Looy, Vereyen & Schmoch (2014). Some IPC classifications could not be automatically linked with appropriate NACE classifications, and hence, these problem cases (4 of them) were removed. These measures produced a list of 84 NACE division, group, and class symbols. Correlations on table 4.4 are calculated from this data set. Lastly, the list of 84 NACE classifications was shortened to 26 classifications by abbreviating the NACE classifications to division level and combining the data elements. Correlations on table 4.5 are calculated from this data set.

For the purpose of the regression analysis, NACE group level data was combined with industry data from the year 2007. However, additional industry data was not available for 8 industries. 7 of these industries had less than 150 combined UM and patent observations, while one industry had 632 observations. Industries that could not be merged with higher hierarchical levels were remitted from the final data[30]. Thus, the final data which was used in regression analysis of chapter 4.3.2 included 76 observations (NACE group level industries) with 11 kinds of variables: UM (firm, individual, and total) counts, patent counts, company counts, employee counts, average turnover, aggregated wages, wage per employee, turnover per company, and turnover per employee.

30 Industries that were removed were NACE classifications 12, 19, 20.2, 25.91 and 20.6, while industries 23.42, 26.52 and 27.33 were merged with industries 23, 26.5 and 27.3

4.3 Empirical findings

4.3.1 Descriptive statistics

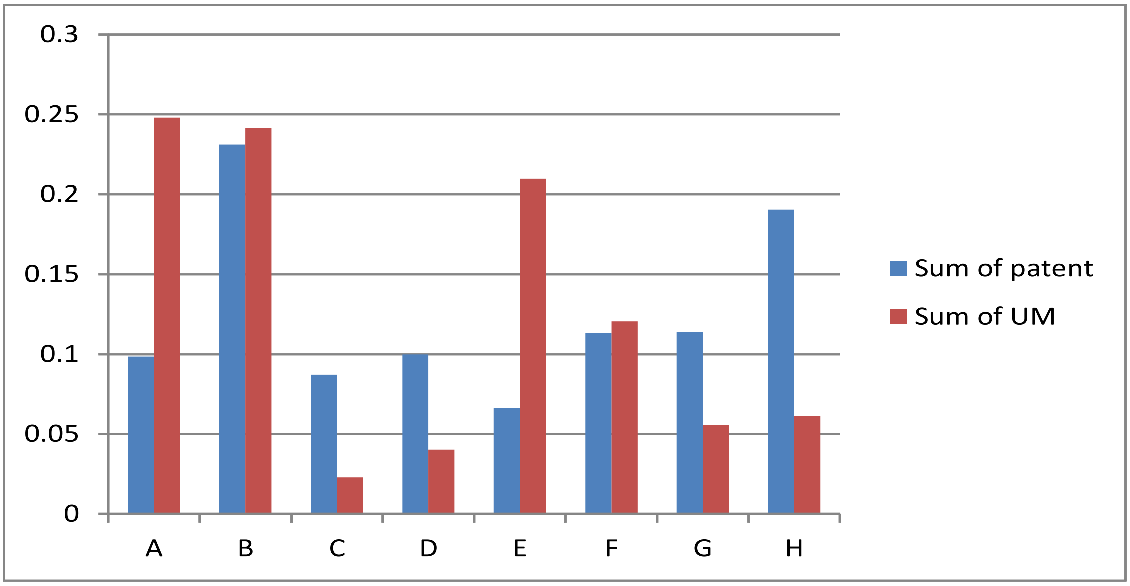
Table 4.1 summarizes key data parameters which include total application counts, the total of IPCs contained in these applications, average IPC count per application, median IPC number per application, maximum and minimum of IPCs in a data set, and the average weight of each IPC. It is possible to make the first observations about the incrementality of utility models at this stage. Firstly, as can be observed from table 4.1, patents are broader than utility models, if the breath or the scope of a patent is measured by the number of IPCs it contains. Secondly, patents include higher hierarchical level classifications, i.e. IPC sub-classes, while utility models only include more specific classifications (IPC main groups and subgroups). Both of these facts are indicative of the incremental nature of inventions protected by utility models.

Table 4.1: Description of data

	Patent	UM	Firm UM	Indiv. UM
Applications	10 815	5 762	3 326	2 436
IPCs	39 841	9 571	5 769	3 802
Average IPC no. per application	3.68	1.66	1.73	1.56
IPC median	3	1	1	1
IPC max. number	86	5	5	5
IPC min. number	1	1	1	1
IPC avg. weight	0.27	0.76	0.74	0.79

First of all, figure 4.3 illustrates the proportional spread of utility model and patent IPC sections that were listed in chapter 4.2. Over 60% of utility model IPCs belong to three sections: human necessities (A), performing operations (B), and fixed constructions (E). Furthermore, relatively few utility models include electricity (H), chemistry (C), textiles and paper (D), and physics (G). However, these allocations are only directional since most of the information contained in an IPC resides in sub-class symbols.

Figure 4.3: Share of IPC sections of total UM and patent IPCs



Continuing, as was imparted in the previous chapter, table 4.2 indicates the correlation results form IPC group level data. These results reveal that the technological content of patents and utility models is very different (correlation of 0.0758), while the technological content of utility models filed by firms and individuals also relatively different (correlation of 0.2762). Furthermore, the technological content of firm utility models seems to be closer to that of patents (correlation of 0.1264 against that of -0.0196).

Table 4.2: IPC group level correlations (16 986 IPCs)

	UM	UM(company)	UM(individual)	Patent
UM	1.0000			
UM(company)	0.8416	1.0000		
UM(individual)	0.7515	0.2762	1.0000	
Patent	0.0758	0.1264	-0.0196	1.000

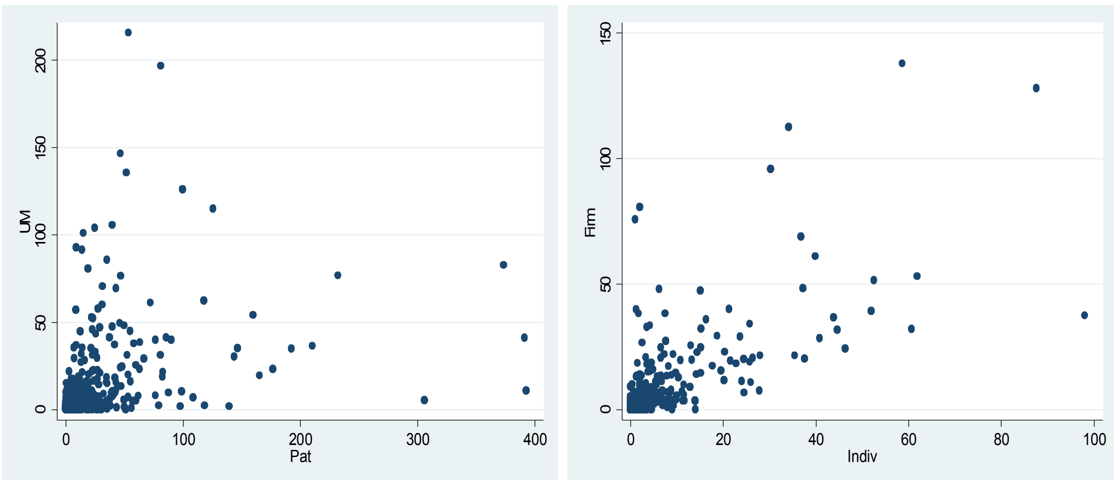
Table 4.3 lists correlations that were calculated from abbreviated IPC subclass data. On this hierarchical level correlations increase across the board, but the differences in technological content remain. The technological content of patents and utility models correlates slightly (0.3744), yet more with firm utility models (0.4735). On the other hand, the technological content of utility models filed by firms and individuals seems to be relatively similar (0.7131) on this stature level of IPC.

Table 4.3: IPC subclass correlations (575 IPCs)

	UM	UM(company)	UM(individual)	Patent
UM	1.0000			
UM(company)	0.9488	1.0000		
UM(individual)	0.8981	0.7131	1.0000	
Patent	0.3744	0.4735	0.1716	1.0000

Differences in the technological content of utility models and patents are also evident in figure 4.4 which maps the scatter patterns of all utility model and patent IPCs on sub-class level. The plotter on the right produces the correlation of 0.3744 between utility models and patents, while the plotter on the left produces the correlation of 0.7131 between firm and individual utility models.

Figure 4.4: scatter patterns of patents and UMs (right) and firm and individual UMs (left)

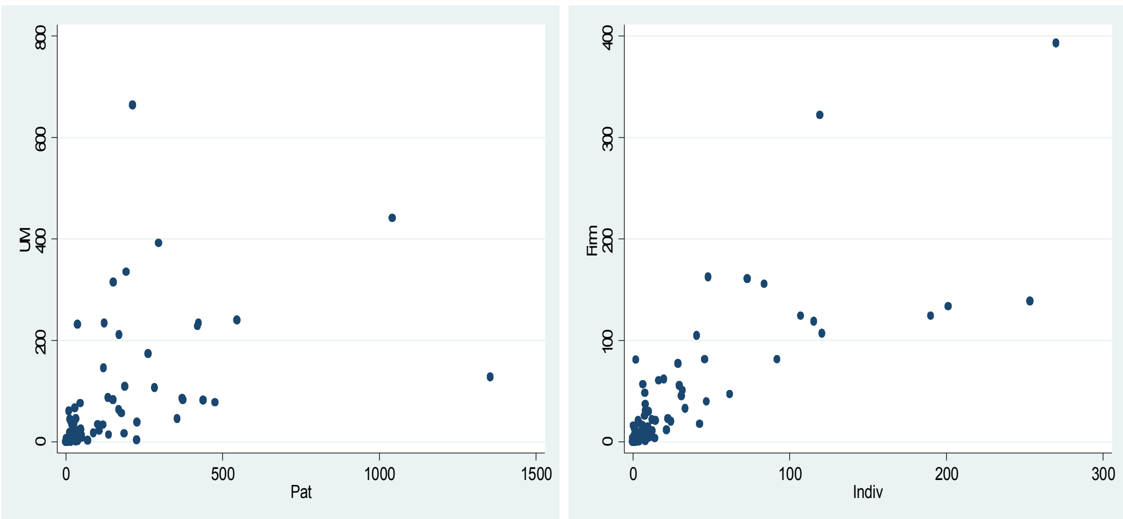


Sub-class level IPCs can be transformed into industrial classifications (NACE2s) according to the concordance scheme developed by Van Looy, Vereyen & Schmoch (2014). Table 4.4 illustrates correlations between utility model and patent NACEs (84 group level observations). Associations of utility models and patents with industries establish a picture that is very similar to what was observed with technological indicators, albeit again with higher correlations. However, as is evident form the left side of figure 4.5, which denotes the scatter pattern of utility models' and patents' industrial associations, the industrial data includes three outlier industries: 26.3 Manufacture of Communication Equipment (most patent), 28.9 Manufacture of Other Special Purpose Machinery, and 43 Specialised Construction Activities (most UM).

Table 4.4: NACE correlations (84 NACEs)

	UM	UM(company)	UM(individual)	Patent
UM	1.0000			
UM(company)	0.9623	1.0000		
UM(individual)	0.9421	0.8154	1.0000	
Patent	0.4992	0.5650	0.3660	1.0000

Figure 4.5: Scatter-plot diagram of UM and patent counts (left) and firm andindividual UM counts (right)



The right hand side of figure 4.5 indicates dispersion firm and individual UMs which produces the correlation 0.8154. The left hand side of diagram 4.5 produces the correlation 0.4992.

In a similar manner, figure 4.6 illustrates the spread of utility model and patent IPC-industry associations for industries that had more than a hundred total connections. Again, Finnish patenting seems to be concentrated in two industries: NACE groups 26.3 and 28.9. The use of utility models, on the other hand, seem to be more spread out across industries, while industry 43 seems to be associated with utility models relatively frequently.

Figure 4.6: NACE spread of IPCs

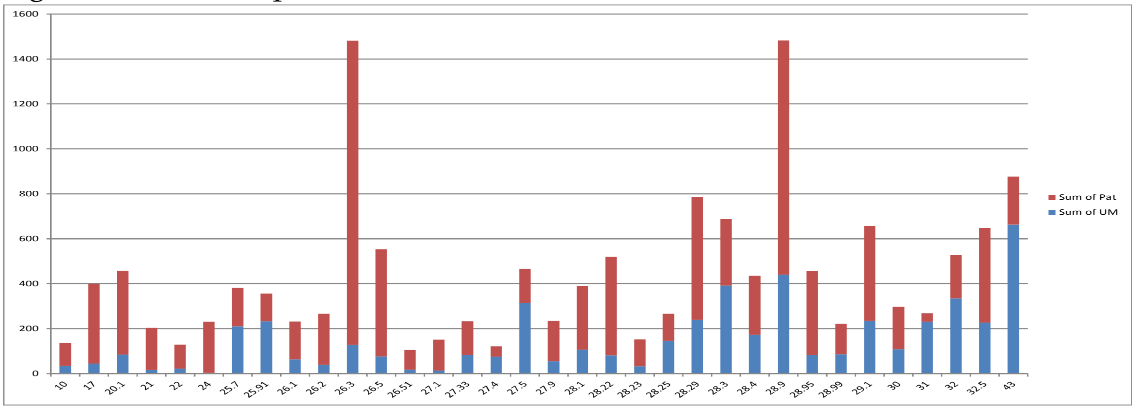


Table 4.6 lists the industries where UMs made up the bulk of all industry connections. These industries also constituted 51% of all UM counts and 14% of patent counts. While the overall number of patents and utility models varies between the listed industries, it is still possible to infer something about the nature of utility models from this data. Many of the industries listed on table 4.6 (e.g. manufacture of furniture, specialised construction activities, and some extent manufacture of wearing apparel) belong to the supplier-dominated industries (Pavitt, 1984), which are characterized as industries where the average firm size is small and the inventive activity is concentrated around making small improvements to technologies that originate in other sectors. However, many other industries (e.g. manufacture of agricultural and forestry machinery and manufacture of cutlery, tools and general hardware) belong under the umbrella of scale-intensive industries. Hence, it is not entirely possible to link incremental innovations with any single sector of the economy denoted by Pavitt's (1984) taxonomy.

Table 4.6: Industries where UM's constituted more than half of all patented inventions

NACE	Pat	UM	UM/patent %	Description
14	9.75	60.08	616 %	Manufacture of Wearing Apparel
31	37.59	231.12	615 %	Manufacture of Furniture
28.14	12.34	44.53	361 %	Manufacture of General Purpose Machinery
43	213.86	663.15	310 %	Specialised Construction Activities
42.2	16.35	39.85	244 %	Construction of Utility Projects
32.9	29.24	66.33	227 %	Manufacturing N.E.C.
27.5	151.47	314.23	207 %	Manufacture of Domestic Appliances
25.91	122.61	233.68	191 %	Manufacture of steel drums and similar containers
32	192.79	334.80	174 %	Other Manufacturing
15	19.46	32.95	169 %	Manufacture of Leather and Related Products
27.4	46.13	75.90	165 %	Manufacture of Electric Lighting Equipment
23.42	11.67	17.50	150 %	Manufacture of Ceramic Sanitary Fixtures
28.92	32.72	44.98	138 %	Manufacture of Machinery for Mining, Quarrying and Construction
25.9	26.41	35.32	134 %	Manufacture of other fabricated metal products
28.3	296.10	391.83	132 %	Manufacture of agricultural and forestry machinery
25.7	170.32	210.32	123 %	Manufacture of cutlery, tools and general hardware
28.25	120.96	145.32	120 %	Manufacture of non-domestic cooling and ventilation equipment

4.3.2 Industrial characteristics and utility models

Table 4.7 lists correlation analysis results that were derived from the NACE group level data which was combined with industrial characteristics data (more in chapter 4.3). Correlations which are marked with grey are not significantly different from zero at a 95% confidence interval. The kinds of industries that patent seem to be rather different from those that choose the utility model: Firstly, utility models seem to correlate with the number of firms in a given industry. Hence, they also correlate with employee numbers and wages (company numbers correlate strongly with employee numbers, 0.8156, and employee numbers with wages, 0,9532). Secondly, industry's turnover does not correlate with utility model counts, even though turnover correlates with wages (0.7586) and employees (0.6143). The same applies to both kinds of utility models, yet, to a smaller extent in the case of utility models filed by individuals.

Patent counts, on the other hand, correlate with entirely different kinds of industries: Firstly, patent counts seem to correlate with average turnover but not with company numbers. Secondly, patents also correlate with such indicators as average wage, average turnover, and turnover per employee. Thus, it can be concluded that some large and productive companies have patented extensively during the period 2000-2013.

Table 4.7: Correlations of NACE group level UM and patent counts (76 observations[31])

	UM	UM company	UM individual	Patent
Companies	0.5628	0.5740	0.4893	0.0147
Employees	0.4331	0.4713	0.3403	0.2660
Turnover	0.1652	0.2011	0.1035	0.6639
Wages	0.3360	0.3844	0.2408	0.4031
Average wage	-0.0677	-0.0046	-0.1382	0.5134
Average turnover	0.0130	0.0330	-0.0130	0.6478
Turnover/employee	0.0124	0.0319	-0.0129	0.6309

Following regression analysis (results are summarized below on table 4.8) hopes to explain the observed industrial differences in the choice of employed appropriability mechanism. Hence, the dependent variable will be the

31 As is stated in chapter 4.3, some NACE group level observations had to be remitted doe to missing industry data.

industrial utility model count and the explanatory variables will be the aforementioned industry characteristics and observed patent counts. All regressions use heteroscedasticity-consistent standard errors. Standard errors are listed below the coefficients and stars denote significance levels.

The first regression consists of the dependent variable and one explanatory variable: patent count. It is observed that patent counts have a

positive effect on observed utility model counts, even on 1% significance level. However, patent count regressor alone explains relatively little about the observed variation as the data included three outlier industries. The second regression adds firm count as an explanatory variable. The second explanatory variable is also observed to have a significant positive effect on observed utility model counts on 1% level. However, the constant is now significant only on 10% level. Yet, the coefficient of determination of the second regression tells that the this model explains over half of the variability of the response data around its mean.

The third regression adds three more explanatory variables: employee number, aggregated turnover and wages. In addition to patent counts and firm counts, these three variables are also observed to be significant on 1% level. The constant is also observed to be significant on 5% level. Industry's aggregated turnover and aggregated wages seem to have a negative effect on counted utility model connections, while employee and firm numbers have a positive effect.

The fourth regression switches three regressors – the employee number, turnover, and wage – to three other regresors: average wage, average turnover, and turnover per employee. Average wage and average turnover are observed to have significant negative effect on utility model counts even on 1% level while turnover per employee is not significant even on 10% level. Patent and firm counts remain highly significant. However, this change of explanatory variables does not seem to increase the coefficient of determination that much.

The fifth regression supplants turnover per employee regressor with employee number. Aggregated turnover and wage regessors are left out as average wage and average turnover are linear functions of other regressors. All regressors and the constant are significant on 1% level. As before, industry's patent and firm counts have a positive effect on observed utility model counts, while employee number, average wage, and average turnover have a negative effect on utility model counts. Moreover, this regression explains over 72% of the observed variation.

Table 4.8 Results of regressions of UM count on the patent count and industry characteristics

Dependent variable: UM count

Regressor	(1)	(2)	(3)	(4)	(5)
Patent count	0.261**	0.257**	0.456**	0.453**	0.472**
	(0.098)	(0.092)	(0.047)	(0.041)	(0.044)
Firm #		0.026**	0.024**	0.023**	0.035**
		(0.002)	(0.004)	(0.001)	(0.003)
Employee #			0.007**		-0.003**
			(0.002)		(0.0008)
Turnover			-0.00001**		
			(1.86e-06)		
Wages			-0.0002**		
			(0.00005)		
Average wage				-6.15**	-4.276**
				(2.128)	(1.571)
Average Turnover				-0.0007**	-0.0006**
				(0.0002)	(0.00009)
Turnover per employee				0.090	
				(0.111)	
Intercept	38.468**	22.089	18.446*	214.615**	177.423**
	(13.71)	(11.663)	(7.602)	(67.004)	(59.395)
Summary statistics					
root MSE	105.39	81.981	67.882	67.348	64.964
R^2	0.2337	0.5426	0.6993	0.7040	0.7246
n	76	76	76	76	76

* = 5% ** = 1%

The fifth regression was then replicated with firm and individual utility models. All regressors remained significant on 1% level, except for average wage which was significant only on 5% level, in both regressions. Coefficients also had the same effects on utility model counts as in regression five of table 4.8. The Coefficients of determination were 0.8029 for firm utility model regression and 0.5137 for individual utility model regression. Once again individuals seem to diverge form firms in their behavior. The signal of individual UM inventions seems to be more “erratic” when compared with firm utility models.

Overall, the picture that emerges from the correlation and regression analyses is one where the dominant factor behind the choice between a patent and an utility model might be *the cost* of this appropriability mechanism, as well as, the lower patentability requirements associated with it. Industries that utilize UMs seem to be those where the firms are numerous and and the average turnover is small. Relatedly, these industries also have low wages. Patents, on the other hand, characterize industries with high aggregate turnover, high turnover per firm, and high turnover per employee. Subsequently, wages are relatively high in these industries.

As it is entirely possible that some patented inventions are also incremental inventions, it can be that industries that have looser financial constraints , i.e. larger margins, are more susceptible to choosing a patent over an utility model[32]. However, this behavior is always restricted by the higher patentability requirements of patents. Hence, as is already stated, utility models are always incremental inventions but the incrementality of patents depends on the financial leeway of firms, and subsequently, industries.

These differences in the industry specific use and importance of utility models demands that the focus of the study of incrementality must be concentrated on a specific industry, and on one that produces mechanical devices. This is because of machinery is by definition made of a system that embodies a base principle of operation and is made of a number of subsystems. Moreover, focusing on the incrementality of a single industry enables the study of utility model's and patent's hierarchical system positions, i.e. how incremental are incremental inventions measured by utility models? Appendix C will shed light on these matters with a case study of incrementality in Finnish paper-making machine industry.

32 Firms in these industreis might also use both forms of legal protection, for example because utility models can be recieved faster.

4.3 Interpretation of empirical findings

Altogether, incremental inventions, measured by utility models, seem to take place in most industries that patent (see figure 4.6). However, the proportion of these inventions varies according to the nature and maturity of industry in question. The industry specific use of utility models can be explained by following industry characteristics: patenting frequency, firm count, employee number, average wage, and average turnover, where the first two characteristics have a positive effect and the latter three have a negative effect on industry's utility model counts. It is evident from the correlation and regression analyses of the previous chapter that Finnish utility models can not be associated with industries with particularly high turnover. Hence, it would seem that incremental inventions are concentrated in industries with many small firms – such as the specialized construction industry.

However, it is also evident that most industries use a variety of different technologies, and hence, an industry's utility model usage tells about the relative position of a connected group of technologies, some of which reside in the era of ferment and some of which reside in the era of incremental change. A closer analysis of incrementality must also encompass technological aspects of utility models. It is evident that patents are broader than utility models (max. 86 patent IPCs in the sample against max. 5 utility model IPCs). This alone is indicative of utility models' incremental nature. At the IPC group level, the technological content of patents and utility models, measured by IPC counts does not produce any remarkable correlations. However, technological content of firm utility models seems to be more like that of patents' while utility models filed by individuals seems to be very different from that of patents. These correlations, however, increase when IPCs are abbreviated to subclass level.

Overall, utility models can be thought of as incremental inventions, and these inventions are indicative of slower technological development in certain industries (roughly all industries in the sample interval except for communication equipment industry). Nevertheless, as the approach chosen for this study is such a novel one, perhaps the largest contribution of this empirical section is the demonstration of this method's functionality in the study of technological change and innovation typology.

5 CONCLUSIONS

5.1 Discussion of key findings

From the literature review that was presented in chapter 2.2 it is apparent that no other study has expressively concentrated on incremental invention and innovation before this one. This might be because of a lack of theoretical understanding of what incremental innovation actually is and how it could be recognized. This study has tried to resolve both of the aforementioned conundrums: To establish a coherent theory around innovation typology, and to measure the industrial salience and the technological content of these kinds of inventions.

The theory behind incremental invention rests on three concepts: base principle (Arthur, 2007), system hierarchy (Arthur, 2007; Murmann & Frenken, 2006), and the dominant design (Murmann & Frenken, 2006). Incremental invention seems to be one that does not introduce a new base principle, i.e. it is not recombination of ideas (Weitzman, 1998) but a development of an existing idea. Further, incremental invention requires some kind of structure or unit of measure within a particular technology. Such a structure can be found in the interrelations or linkages of individual components and subsystems (Murmann & Frenken, 2006). In this view, incremental inventions apply to peripheral components in terms of new knowledge that they bring to bear (Murmann & Frenken, 2006). Relatedly, as incremental innovations take place in peripheral components, the number of these components destines the possibilities for improvement: The more complex a product is, the larger the possibility to improve it incrementally (Murmann & Frenken, 2006). Lastly, the emergence of a dominant design within a product category ensures that the aforementioned requirements are filled and further developments will take place in peripheral components.

Overall, incremental innovations seem to have a special place in the gestalt of a generic technology's life cycle where the relative salience of incremental innovations tells that a particular technology has reached a certain kind of maturity. Convenient solutions have been found to design problems higher up the system hierarchy, and thus, future problems will be found and solved in peripheral parts of the system. In practice, this often means the emergence of a dominant design, and it is this concept that allows us to tie the discussion around incremental innovation to larger context.

The emergence of a dominant design can be seen as a watershed moment in an industry's evolution. According to the theoretical findings of this study, what seems to be happening is that the R&D procedure itself changes its form from stochastic into deterministic one with the emergence of a dominant design. The logic behind of this argument should be evident from the previous discussion: A dominant design essentially means the solidification of core components, and hence, further developments can be said to be the refinement of an existing design. What is important here is that the design already *works*. The same can not be said about design trials that take place before this watershed moment.

This changes the relative incentives to partake a patent race, i.e. the game that is played changes from Gilbert and Newbery's (1982) one to Reinganum's (1983) one. Essentially, this means that incremental innovations can be associated with established firms and mature industries. Established firms have an incentive and the capability (Henderson, 1993) to protect their position in the market with improvements to their products, and these improvements are likely to be small (Reinganum, 1983), and this behavior is largely incentivized by the existence of patents (Gilbert & Newbery, 1982). If, then, incremental innovations follow the emergence of a dominant design, they are likely to be process innovations (Klepper, 1997). Essentially, this means a shift from product characteristics (quality) to production costs (productivity).

All this involves a remarkable similarity with the observed "inverted-U" pattern of entry and exist in some industries (Murmann & Frenken, 2006). In the era of ferment (prior to the emergence of a dominant design) market leadership changes hands frequently and competition is fierce – more and more firms enter the marketplace. Then, some designs prove better than others or some designs are selected because of network effects to be better than others. Firms that have invented these designs see their market shares rise and the market becomes more oligopolistic. If imitation is possible and one design rules them all, a dominant design is likely to emerge. After this threshold, innovations are likely to be incremental. Therefore the following statement: Technological change and market structure seems to be fundamentally and causally connected.

Theoretical findings of this study can also be thought of as a form of aggregated *learning* where a given technology's – base principle's – full potential is achieved, i.e. as the *learning curve*. Invention, on the other hand, can be though of as birth of a new technology, i.e. radical invention. As it so happens, both of these phenomena have been modeled in growth literature. Young (1993) and Stein (1997) present combinations of the aforementioned models. Young's (1993) model can be used to model the development of whole economies and Stein's (1997) model is more suitable for the examination of industry evolution. More on these models in appendix B.

Yet, how can these developments in the product or the underlying technology be measured? By reviewing the prior empirical literature it was discovered that perhaps the simplest solution to this conundrum is the use of patent statistics. Utility models, especially, offer a novel method for the study of technological change. A growing evidence points to the fact that a growing number of patent applications might embody small or even trivial inventions (Guellec & van Pottelsberghe de la Potterie, 2000), while only a small percentage of patents prove to be valuable (Scherer & Harhoff 2000; Silverberg & Verspagen, 2007). Utility models, on the other hand, are always likely to be small and incremental inventions.

Using this method it was found that technologies and industries that could be said to be in the era of incremental change, measured by the relative prevalence of utility models, were mostly supplier-dominated industries (Pavitt, 1984). However, it is also apparent that those industries that patent also use utility models – to a varying degree. Furthermore, it is clear that these results are at least to some extent driven by selection. A large firm that innovates incrementally always has an incentive to protect its invention as broadly as possible as it can spread the cost of doing so over many products. A small firm does not have this option, and hence, such a firm should be more careful in choosing the form of legal protection.

5.2 Implications

Understanding the dynamics related to the evolution of a generic technology is of utmost importance for designers of an industrial policy – if such a policy is ever to come into being. In a world of low trade barriers, industries can be thought of as regional or even global. Hence, technological dynamics illustrated in this study can have tremendous effects on national industrial dynamics. Hence, an industrial policy should have two goals: the generation of novel ideas[33] and propelling these ideas towards a dominant design.

The theoretical and empirical findings of this study have two kinds of implications for future research, accordingly. Starting with the theoretical findings, a deeper understanding of technologies' evolution should be incorporated in models of growth and industry evolution. This is because different rules seem to apply in the two eras of technologies' evolution: the era of ferment is likely to see a patent races à la Reinganum (1983) while the era of incremental change is likely to see races à la Gilbert and Newbery (1982). Continuing with the empirical methodology of this study, it is apparent that further research that utilizes the method of this study should examine the temporal changes in the relative importance of utility models and patents. According to the theoretical findings presented above time and time again, the share of utility models (incremental inventions) should increase as a technology, or even an industry, matures. This, a topic for another and larger study. Future research should also control for any selection that might arise from the relative sizes and turnovers of individual firms. Thus, it might be appropriate to center around some specific industry.

All in all, it needs to be concluded that understanding technological dimensions of economic matters should be part of any economist's training. Technological dynamics and industrial structures might indeed be causally connected.

---

33 In practice, this means creation of new "needs" that do not exist in a market economy. Remember that a radical invention is in practice an idea that connects an need with an effect to satisfy it (Arthur, 2007). Normal human needs can be characterized as finite and classifiable, while state needs can be thought of as a different and, perhaps, a larger set. Think of colonialism, conflict, and space exploration. Expressed in a simple manner: market making is an important function for the state.

Bibliography

Abernathy, W. J. & Utterback, J. M. 1978. Patterns of industrial innovation. Technology review 64, 254-228.

Aghion, P., Harris, C., Howitt, P. & Vickers, J. 2001. Competition, imitation and growth with step-by-step innovation. The Review of Economic Studies 68 (3), 467-492.

Aghion, P. & Howitt, P. 1992. A Model of Growth Through Creative Destruction. Econometrica 60 (2), 323-351.

Anderson, P. & Tushman, M. L. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. Administrative Science Quarterly , 604-633.

Anton, J. J. & Yao, D. A. 2004. Little patents and big secrets: managing intellectual property. RAND Journal of Economics , 1-22.

Arrow, K. 1962. Economic welfare and the allocation of resources for invention. In The rate and direction of inventive activity: Economic and social factors. Princeton University Press, 609-626.

Arrow, K. J. 1962. The Economic Implications of Learning by Doing. The Review of Economic Studies 29 (3), 155-173.

Arthur, W. B. 2007. The structure of invention. Research Policy 36 (2), 274-287.

Atal, V. & Bar T. 2014. Patent Quality and a Two-Tiered Patent System.The Journal of Industrial Economics 62 (3): 503-540.

Beneito, P. 2006. The innovative performance of in-house and contracted R&D in terms of patents and utility models. Research Policy 35 (4), 502-517.

Carlino, G. & Kerr, W. R. 2014. Agglomeration and innovation .

Chang, H. F. 1995. Patent scope, antitrust policy, and cumulative innovation. The Rand journal of economics , 34-57.

Cohen, W. M., Nelson, R. R. & Walsh, J. P. 2000. Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not) .

Dahlin, K. B. & Behrens, D. M. 2005. When is an invention really radical?: Defining and measuring technological radicalness. Research Policy 34 (5), 717-737.

Dosi, G. 1982. Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. Research policy 11 (3), 147-162.

Dosi, G. & Nelson, R. R. 2013. The evolution of technologies: an assessment of the state-of-the-art. Eurasian Business Review 3 (1), 3-46.

Freeman, C. 1995. The ‘National System of Innovation’in historical perspective. Cambridge Journal of economics 19 (1), 5-24.

Gilbert, R. J. & Newbery, D. M. 1984. Uncertain innovation and the persistence of monopoly: Comment. The American Economic Review , 238-242.

Gilbert, R. J. & Newbery, D. M. 1982. Preemptive patenting and the persistence of monopoly. The American Economic Review , 514-526.

Green, J. R. & Scotchmer, S. 1995. On the division of profit in sequential innovation. The Rand journal of economics , 20-33.

Greenhalgh, C. & Rogers, M. 2010. Innovation, intellectual property, and economic growth. Princeton University Press.

Griliches, Z. 1990. Patent statistics as economic indicators: a survey.

Gruber, M., Harhoff, D. & Hoisl, K. 2013. Knowledge recombination across technological boundaries: scientists vs. engineers. Management Science 59 (4), 837-851.

Guellec, D. & de la Potterie, Bruno van Pottelsberghe 2000. Applications, grants and the value of patent. Economics letters 69 (1), 109-114.

Hall, B. H. 2004. Exploring the patent explosion. The Journal of Technology Transfer 30 (1-2), 35-48.

Hall, B. H. & Harhoff, D. 2012. Recent research on the economics of patents .

Hall, B., Helmers, C., Rogers, M. & Sena, V. 2014. The choice between formal and informal intellectual property: a review. Journal of Economic Literature 52 (2), 375-423.

Henderson, R. 1993. Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. The Rand journal of economics , 248-270.

Henderson, R. M. & Clark, K. B. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. Administrative Science Quarterly , 9-30.

Klepper, S. 1997. Industry life cycles. Industrial and corporate change 6, 145-182.

Kortum, S. & Lerner, J. 1999. What is behind the recent surge in patenting? Research policy 28 (1), 1-22.

Kuhn, T. S. 1962. The Structure of Scientific Revolutions Vol. .

Kultti, K., Takalo, T. & Toikka, J. 2007. Secrecy versus patenting. The Rand journal of economics 38 (1), 22-42.

Lerner, J. 1994. The importance of patent scope: an empirical analysis. The Rand journal of economics , 319-333.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R. & Griliches, Z. 1987. Appropriating the returns from industrial research and development. Brookings papers on economic activity 1987 (3), 783-831.

Mansfield, E. 1986. Patents and innovation: an empirical study. Management science 32 (2), 173-181.

Moser, P. 2003. How do patent laws influence innovation? Evidence from nineteenth-century world fairs .

Mowery, D. & Rosenberg, N. 1979. The influence of market demand upon innovation: a critical review of some recent empirical studies. Research Policy 8 (2), 102-153.

Murmann, J. P. & Frenken, K. 2006. Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. Research Policy 35 (7), 925-952.

Nelson, R. R. & Winter, S. G. 1977. In search of a useful theory of innovation. In Innovation, Economic Change and Technology Policies. Springer, 215-245.

Nikulainen, T. 2008. A historical perspective on patents and industry growth in Finland: Connections and lags. No. 1159. ETLA discussion paper.

Nordhaus, W. D. & Nordhaus, W. D. 1969. Invention, growth, and welfare: A theoretical treatment of technological change. MIT press Cambridge, MA.

O'Donoghue, T. 1998. A patentability requirement for sequential innovation. The Rand journal of economics , 654-679.

O'donoghue, T., Scotchmer, S. & Thisse, J. 1998. Patent breadth, patent life, and the pace of technological progress. Journal of Economics & Management Strategy 7 (1), 1-32.

O'Donoghue, T. & Zweimüller, J. 2004. Patents in a model of endogenous growth. Journal of Economic Growth 9 (1), 81-123.

Pavitt, K. 1984. Sectoral patterns of technical change: towards a taxonomy and a theory. Research policy 13 (6), 343-373.

Porter, M. E. 1980. Competitive strategy: techniques for analyzing industries and competitors.

Reinganum, J. F. 1989. The timing of innovation: Research, development, and diffusion. Handbook of industrial organization 1, 849-908.

Reinganum, J. F. 1985. Innovation and industry evolution. The Quarterly Journal of Economics , 81-99.

Reinganum, J. F. 1984 . Practical implications of game theoretic models of R&D. The American Economic Review 74 (2), 61-66.

Reinganum, J. F. 1984. Uncertain innovation and the persistence of monopoly: Reply. The

American Economic Review 74 (1), 243-246.

Reinganum, J. F. 1983. Uncertain innovation and the persistence of monopoly. The American Economic Review 73 (4), 741-748.

Romer, P. M. 1990. Endogenous Technological Change. Journal of Political Economy 98 (5), S71-S102.

Rosenberg, N. 1982. Learning by using. Inside the black box: technology and economics , 120-140.

Rosenberg, N. 1976. On technological expectations. The Economic Journal 86 (343), 523-535.

Rosenkopf, L. & Nerkar, A. 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. Strategic Management Journal 22 (4): 287-306.

Sahal, D. 1985. Technological guideposts and innovation avenues. Research policy 14 (2), 61-82.

Scherer, F. & Harhoff, D. 2000. Policy implications for a world with skew-distributed returns to innovation. Research Policy 29 (4), 559-566.

Schmoch, U., Laville, F., Patel, P. & Frietsch, R. 2003. Linking technology areas to industrial sectors. Final Report to the European Commission, DG Research 1 (0), 100.

Schumpeter, J. A. 1942. Socialism, capitalism and democracy. Harper and Brothers.

Schumpeter, J. A. 1939. Business cycles. Cambridge Univ Press.

Scotchmer, S. 1991. Standing on the shoulders of giants: cumulative research and the patent law. The Journal of Economic Perspectives 5 (1), 29-41.

Scotchmer, S. & Green, J. 1990. Novelty and disclosure in patent law. The Rand journal of economics , 131-146.

Shane, S. 2001. Technology regimes and new firm formation. Management science. 47 (9): 1173-1190.

Silverberg, G. & Verspagen, B. 2007. The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance. Journal of Econometrics 139 (2), 318-339.

Smith, A. 1776. An inquiry into the nature and causes ofthe wealth ofnations. London: George Routledge and Sons .

Stein, J. C. 1997. Waves of creative destruction: Firm-specific learning-by-doing and the dynamics of innovation. The Review of Economic Studies 64 (2), 265-288.

Stuart, T. E., & Podolny J. M. 1996. Local search and the evolution of technological capabilities. Strategic Management Journal. 17 (S1) : 21-38.

Tirole, J. 1988. The theory of industrial organization. MIT press.

Tushman, M. L. & Anderson, P. 1986. Technological discontinuities and organizational environments. Administrative Science Quarterly , 439-465.

Usher, A. P. 2013. A History of Mechanical Inventions: Revised Edition. Courier Corporation.

Utterback, J. M. & Suarez, F. F. 1993. Innovation, competition, and industry structure. Research policy 22 (1), 1-21.

Van Zeebroeck, N. 2011. The puzzle of patent value indicators. Economics of Innovation and New Technology 20 (1), 33-62.

Verhoeven, D., Bakker, J. & Veugelers, R. 2016. Measuring technological novelty with patent-based indicators.Research Policy 45 (3), 707-723.

Weitzman, M. L. 1998. Recombinant growth. Quarterly Journal of Economics , 331-360.

Young, A. 1993. Substitution and complementarity in endogenous innovation .

Young, A. 1991. Invention and bounded learning by doing .

Zuscovitch, E. 1986. The economic dynamics of technologies development. Research Policy 15 (4), 175-186.

Online citations:

WIPO (World Intellectual Property Organization). 2015. Cited 1.10.2015:
http://www.wipo.int/sme/en/ip_business/utility_models/utility_models.htm
OEC (Observatory of economic complexity). 2016. Cited 10.4.2016:
http://atlas.media.mit.edu/en/profile/country/fin/
NACE (Rev. 2, 2008). Cited 26.4.2016:
http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF
PRH (Patentti ja rekisterihallitus). 2015. Cited 1.10.2015:
https://www.prh.fi/stc/attachments/patentinliitteet/hmopas.pdf
STATFI (Statistics Finland). 2016. Cited 10.4.2016:
http://www.stat.fi/til/syr/2007/02/syr_2007_02_2009-02-10_tie_001.html
IPC(version 2015). Cited 14.5.2016:
http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf

Appendix A - Incremental innovations and intellectual property rights

In the theoretical patent literature, innovation size can be thought of as a function of market structure, or it can be thought of as a function of the breadth of the exclusion right that a patent provides[34]. Yet, it is obvious that technological change is affected by *both* kinds of the aforementioned incentives. Hence, competition and patent policy might have complementary roles in achieving optimal incentives for innovation (Aghion et. al., 2001). Understanding this is crucial for understanding incremental innovation. Chapter 2.3 covered the market structure related incentives and this appendix will deal with patent related incentives.

An inventor is granted a temporary monopoly right in the production of his invention in exchange of early publication of the invention and in order to create an incentive for him to invent in the first place (Hall & Harhoff, 2012). The early publication of research, serves the purpose of forestalling duplicative research, i.e. removing the need to "reinvent the wheel". The right to exclude others from producing the invention guarantees a potential monopoly and enables the commercialization of ideas. Yet, it is also worthwhile to remember that the incentive for innovation exists irrespective of intellectual protection rights (Arrow, 1962b). These two justifications for patents are labeled *disclosure* and *appropriability*, respectively. With obvious overlap, these justifications divide the economic patent literature into two broad halves: one examines the choice between patenting and secrecy and the other focus on optimal *patent-design*.

When an innovator chooses patenting, he relinquishes his option to protect the innovation with secrecy. Other appropriation methods include secrecy, lead-time, confidentiality agreements, and complexity (Hall, Helmers, Rogers & Sena, 2014). As the patent-design literature emphasizes appropriability, the focus of secrecy related literature is *disclosure*, i.e. the R&D efficiency. In this view, secrecy means inefficient behaviour because secret knowledge cannot be built upon or utilized in a wider context. Yet, survey evidence suggests that in some cases secrecy offers better protection against imitation (see e.g. Cohen, Nelson & Walsh, 2000). Moreover, the signaling

34 It should be noted that these two incentives have been strikingly separate in economic literature (see e.g. Reinganum, 1983b vs. O'Donoghue, 1998).

model of Anton and Yao (2004) suggest that small process innovations include no incentive for imitation and are thus always fully disclosed, while radical innovations are mainly kept secret. Overall, the strength of the intellectual property rights seems to have less effect on the patenting rate of incremental innovations (Anton & Yao, 2004; Kultti, Takalo & Toikka, 2007).

Next, the context of *appropriability*. The patent-design literature has been shaped by three major ideas: Starting with Nordhaus (1969), innovation was examined in *isolation*. The focus was in the trade-off between creating a monopoly and incentivizing innovation. The second idea was that the *cumulative* nature of technological change brings about situations where innovations are improved, or built on, by secondary innovations (Scotchmer, 1991). Here the underlying premise was in creating an incentive for *both* inventors in a case where the profit from the second innovation should incentivize two innovators – the original and the subsequent one (Scotchmer, 1991; Greene & Scotchmer, 1995; Chang, 1995). The third idea was that instead of having a two-stage innovation process, as was elaborated above, technological change consists of a long *sequence* of innovators, each improving the the underlying technology (O'Donaghue et al., 1998; O'Donoghue, 1998).

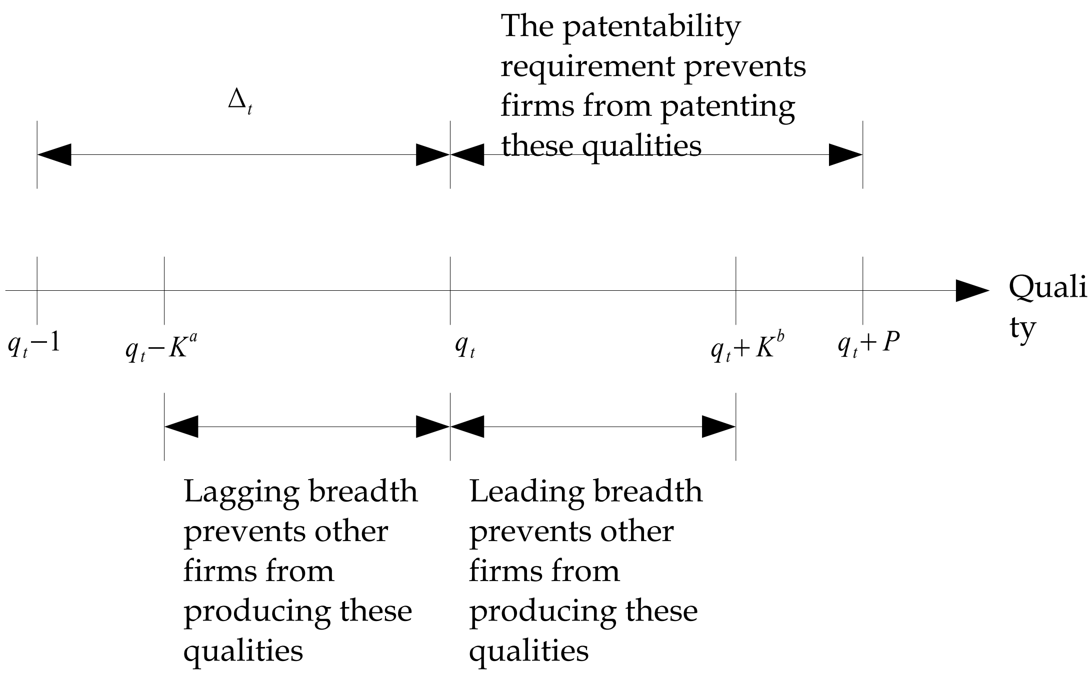
The parts of a patent-design, i.e. the the chosen policy, are patent *breadth*[35], patent *life*, and *patentability requirement* (O'Donoghue and Zweimüller, 2004). Patent breadth can further be divided into *leading* and *lagging* breadth. Lagging breadth protects against imitation from inferior products while leading breadth protects against future, higher quality products. Another part of the patent-design, patent's life, can end either because the patent expires or it may be that another innovation displaces the patented one in the market. The *effective patent life* represents the fulfillment of one of the former prerequisites. Thus, the effective patent life depends on the statutory patent life as well as patent breadth. Yet, a patent is not always granted – the underlying invention is required to have certain characteristics, i.e it needs to fulfill a *patentability requirement*. In Europe, these requirements are *novelty*, *inventive step* and *industrial applicability*. However, these concepts are rather vague in explicit economic meaning since there is no straightforward way of interpreting them.

Figure A1 illustrates these concepts. Δ_t Is the innovation size of quality q_i at time t. K^a portrays lagging breadth and K^b leading breadth, while P

35 Also known as *patent scope* (e.g. Chang, 1995).

is the patentability requirement for innovation t + 1[36]. Because of lagging breadth, no other firm can produce any quality between $q_t - K^a$ and q_t without a license, and because of leading breadth, no other firm can produce any quality between q_t and $q_t + K^b$ without a license. "In other words, rival firms can compete without a license only if they produce a quality $q \leq q_t - K^a$ or a quality $q \geq q_t + K^b$. The patentability requirement specifies the minimum patentable innovation size for the subsequent generation, so no firm can patent any quality between q_t and $q_t + P$ " (p. 661). (O'Donoghue, 1998).

Figure A1. Illustration of patent characteristics.



Adopted from O'Donoghue (1998)

O'Donaghue et al. (1998) propose that lagging breadth alone may provide insufficient incentives for invention even when the statutory patent life is very long[37]. In addition, they show that leading breadth can extend effective patent

36 There is no necessary connection between K^b and P – it is possible for $K^b \leq P$ or $K^b \geq P$ (O'Donoghue, 1998).

37 O'Donaghue et al. (1998) assume that ideas are private information. In contrast, the patent race literature, which was briefly reviewed in chapter 2.3, assumes that potential

life and increase R&D (O'Donaghue, Scotchmer & Thisse, 1998). Yet, it should be noted that O'Donoghue et al. (1998) do not employ innovation typology in their analysis. This is because they assume that the rate and the size of ideas is exogenously determined. However, as remarked by O'Donoghue (1998), firms also choose the size of the innovation that they pursue. The possibility to make this choice is fundamental for the theory surrounding incremental and radical innovations.

O'Donoghue (1998) examined the effect of patentability requirements in an infinite sequence of non-drastic innovations where firms repeatedly supersede each other[38]. This never ending sequence of innovations produces a situation where there is no efficiency effect. However, there can be a replacement effect in the model. In this simplified and hypothetical situation higher patentability requirements can cause firms to pursue larger innovations than they normally would because of longer potential market incumbency and, hence, increased rewards to innovation[39]. Thus, patentability requirements are essentially a form of protection against future innovators. Altogether, patentability requirements can prevent firms from pursuing suboptimally small inventions (La Manna, 1992; Luski and Wettstein, 1995, both cited in O'donoghue, 1998), increase R&D spending (O'donoghue, 1998), and encourage firms to "stay in there race" if they happen to fall behind in it (Scotchmer & Greene, 1990).

Empirical research into patenting has been encouraged by the increase in patenting across the world (see e.g. Kortum & Lerner, 1999), as well as, availability of patent data in a machine readable form. This research has reestablished the long known fact about the skewness of patent value distribution (Griliches, 1990; Scherer & Harhoff 2000; Silverberg & Verspagen, 2007). The skewness might be adhering to a log normal law (Scherer & Harhoff, 2000) or it might be characterized by fat tails (Silverberg & Verspagen, 2007). Nevertheless, it seems that a small minority of innovations yield the majority of innovations' total economic value (Scherer & Harhoff, 2000). By combining these two facts – the known skewness of patent value distribution and a surge in overall patenting – it is evident that a growing number of patents might be filed over trivial or "small" inventions (van Zeebroeck, 2011).

However, innovation takes many forms. Therefore it is appropriate to differentiates between industries. Survey evidence suggests that patenting

R&D opportunities are public knowledge (e.g. Reinganum, 1989).

38 In innovation and technology studies such a situation would be an oxymoron.

39 That is, if larger innovations are harder to achieve (O'Donoghue, 1998).

practices of various industries differ greatly. Levin et al, (1987) discovered that patent protection is important especially in pharmaceutical industry, while in most industries patents are not at the forefront of competition. Elsewhere Mansfield (1986) suggested that the absence of intellectual property rights would have little impact on innovation in majority of firms in most industries. Yet, pharmaceuticals were again an exception. Apart from patents, firms can protect their inventions with exploitation of lead time, moving rapidly down the learning curve, the use of complementary sales and service capabilities and secrecy (Levin et al., 1987). Further, firms can also to employ more than one appropriation method. Levin et al, (1987) found that industries and product and process innovations vary in the effectiveness of different appropriability mechanisms. In addition, more than one method was judged to be effective. However, patents were not assessed to be one of the most important appropriation methods – except in pharmaceutical industry. (Cohen, Nelson & Walsh, 2000).

Cohen, Nelson & Walsh (2000) reproduced Levin et al.'s (1987) survey analysis. They find that patent are seen as the least important method of appropriation in the majority of US manufacturing firms who tend to emphasize secrecy and lead time. Moreover, the survey evidence also reveals patenting to be motivated by reasons other than profiting directly from a patented invention. Other reasons for patenting were reported to be the use of blocking patents, the use of patents in negotiations, and the prevention of law suits (Cohen et al., 2000). Further, Cohen et al. (2000) identify two kinds of industries: "discrete" product industries, such as chemicals, and "complex" product industries, such as telecommunications equipment or semiconductors. "In the former, firms appear to use their patents commonly to block the development of substitutes by rivals, and in the latter, firms are much more likely to use patents to force rivals into negotiations" (Ibid. p.1).

Relatedly, by using data from 19th century world fairs, Moser (2003) found no evidence that the existence of patents increased overall inventive activity, instead it was found that patents seemed to shape the direction of technological change. In the absence of patent legislation, inventive activity congregated into sectors and industries where secrecy was relatively more important (Moser, 2003). Furthermore, Hall (2004) found the much more recent growth in patent quantity in the United States has taken place in all technologies, but not in all industries, and is concentrated in electronics, computing, and scientific instruments industries. This might be because patents are also viewed important if not essential in securing financing for a new venture (Hall, 2004).

To summarize, the empirical patent literature has established several crucial observations: There has been a surge in overall patenting (Hall, 2004), only a small minority of patented inventions turn out to be truly radical ones (Griliches, 1990; Scherer & Harhoff 2000; Silverberg & Verspagen, 2007), patents are more useful method of appropriation in some industries (Levin et al., 1987; Cohen et al., 2000), and patent policy can direct technological change (Moser, 2003). On the other hand, the theoretical patent literature has established the following ideas, innovations are compelled by market structure (Industrial organization economics) and appropriability conditions, as well as, technological possibilities (evolutionary economics). If possible, small inventions are always patented (Anton & Yao, 2004). Protection against future inventions can stimulate aggregate R&D (O'Donoghue (1998); O'Donoghue & Zweimüller, 2004).

The reason why these ideas have been confined to the appendix is because these ideas are not easily connected with the definition of an incremental invention that was proposed in chapter 2.4.

Appendix B – Incremental innovations as a form of learning

Innovation size can also be examined in the context of endogenous growth theory. Although, it can be said that innovation typology is by and large absent from many of these models, where innovation is often considered as an ubiquitous unity (e.g. models of Romer, 1990 and Aghion & Howitt, 1992). Yet, some of the later generation models (e.g. those of Young, 1993b and Stein, 1997) can be helpful in getting to grips with this phenomenon. The premise behind this is the following: incremental innovation can be thought of as being analogous to *learning*[40]. Here, learning is viewed as a process where an original invention is refined and extended, i.e. the refinement of a previous radical innovation by incremental means. The seminal paper of Arrow (1962a) set out the groundwork in this area. According to him, learning is the product of experience and experience can only be accumulated during activity, hence the term: *learning-by-doing* (Arrow, 1962a).

Models such as learning-by-doing (Arrow, 1962a) and learning-by-using (Rosenberg, 1982) have considered the influence of learning for technological change – from opposite ends of the spectrum. These two forms of learning could be observed in the use of utility models by individuals and companies. Nevertheless, since the definition of incremental invention was restricted to a continuum of inventions, it could not be determined that utility models purely indicated incremental inventions. A more appropriate term for utility model inventions would be a "small" invention. However, it is also obvious that learning alone is a poor explanation for every economic circumstance, decision, and opportunity that can be argued to be a cause of technological development. Nonetheless, learning-by-doing can be combined with other, perhaps more fundamental[41], reasons for economic growth: Some concoctions of models of

40 Here it bears to remember what was written about Dosi's (1982) technological paradigms. Paradigms' manifest themselves in the gestalt of possibilities for the refinement of an original invention, the radical kind. Incremental inventions are restricted by these possibilities. Hence, it can be stated that if these possibilities already exist, so to speak, incremental innovations are simply discoveries associated with the continued use or production of the original invention, and hence, learning.

41 With the broadest generalization possible, four pretexts for economic growth can be identified: division of labour (Smith, 1776), entrepreneurship (Schumpeter, 1942), learning-by-doing (Arrow, 1962a), and ideas and knowledge creation (Weitzman, 1998). With all of these being part of a whole.

learning and “new” and “Schumpeterian” endogenous growth models are Young (1993b) and Stein (1997). In these models learning-by-doing has been merged with an endogenous growth model similar to that of Romer (1990): Young (1993b) – and to a model alike to that of Aghion and Howitt (1992): Stein (1997).

According to Young (1993a), innovation has the possibility of both complementing and substituting older technologies, a fact that is often missing from the earliest generation of "Schumpeterian" growth models (e.g. that of Aghion & Howitt, 1992). Thus, Young (1993b) ameliorates the situation with a model that combines a model of invention (similar to that of Romer, 1990) to a model of learning (similar to that of Arrow, 1962a). Models of invention (e.g. Romer, 1990) assume that innovations achieve their full potential at the moment of their introduction (Young, 1993). On the other hand, models of learning-by-doing make the assumption that potential productivity gains from learning are unbounded (Young, 1993).

Young's (1993b) synthesis addresses both of these shortcomings: In the general equilibrium of this model, the rate of learning and invention are connected with each other. Yet, continuous invention is required to prevent stagnation. However, this model is capable of producing multiple equilibria. With small markets, large rates of time preference, or relatively expensive invention there will be no invention and the model behaves like a model of invention: it emphasizes incentives (e.g. patents) that lure resources from production into invention. With large markets or cheap invention the inventive activity is especially profitable and it surpasses economy's capability to learn. Hence, it is the society's (or industry's) learning capability that prescribes the rate of steady state rate of growth. And lastly,

> "*for intermediate parameter values, both invention and learning are important constraints, and policies aimed at either activity will influence the economy's growth rate*" (Young, 1993. p.448).

The assumptions that lead to the aforementioned equilibria are the following. Invention costs resources while productivity gains from learning are serendipitous. This is because resources that are devoted to R&D are assumed to be useful in production[42]. Every new inventor receives an *unending patent* to

42 According to Freeman (1995) it is possible for incremental innovations to originate from the day-to-day production activities of a firm. In this view, small innovations are serendipitous – just like Young (1993b) assumes. On the other hand, as the last chapter

his or her invention and this creates the incentive for invention because it is assumed that *there is no imitation*. Production experience generates knowledge on how to make products more efficiently *and this learning can not be appropriated by any one actor*. This in turn creates incentives to create many different types products. Productivity gains from learning spill over sectorial boundaries symmetrically. Furthermore, potential quality increments from learning that are associated with each good are assumed to be finite. Further, it is assumed that new products are initially inferior to existing products that have achieved their full potential through learning, yet they also have potential to provide more utility in the future as learning takes place. Lastly, the margin of inferiority increases as the newly invented goods move beyond the society's cumulated learning experience. (Young, 1993).

All in all, Young's (1993b) work elucidates the interconnectedness of incentives for innovation, patterns of demand, and patterns of production in setting limits for economic growth[43]. The “Schumpeterian” growth models (e.g. that of Aghion & Howitt, 1992) paint a somewhat simpler picture of a continuum of firms, with each new one replacing an older incumbent in *a single industry*. In this view, the destruction of existing firms' market shares is an essential part of technological change. In it's simplest form, this process of 'creative destruction' relies solely on radical (drastic) innovations. As before, the approach taken here is that incremental innovation can be thought of as a process of learning. Stein (1997) combines these two ideas, learning-by-doing and creative destruction. The result is a model where the incumbent is able to survive longer, or even indefinitely if it is able to amass large enough learning benefits.

Stein (1997) presents two kinds of arguments for the survival of established firms: Firstly, these firms might have a comparative advantage in innovation (Henderson, 1993). Secondly, it is possible that established firms have other competitive weapons such as loyal customers and well developed distribution chains, or more generally “customer bases” (see Porter 1980). Essentially this means that firms compete on two dimensions – with "high-spillover" dimension, i.e. product quality, and in "zero-spillover" dimension, i.e.

---

proposed, they can also be a result of zemblanitious work of R&D departments in established firms.

43 However, according to the seminal work of Weitzman (1998) “...the ultimate limits to growth may lie not so much in our ability to generate new ideas as in our ability to process an abundance of potentially new ideas into usable form (Weitzman, 1998. p 331).

distribution costs. In the former, an innovation ultimately becomes common knowledge and in the latter case there will be no spillovers, at all. Ultimately this leads to a situation where the timing of innovation is not constant over time: When an incumbent is replaced, and consequently, its customer base is destroyed, a subsequent replacement becomes more likely. This shakeup externality means that innovations happen in waves. Furthermore,

> "*...customer bases can in some circumstances – though this need not always be true—dramatically reduce the long-run average level of innovation. Or stated somewhat more generally, the potential for firm- and product-specific learning- by-doing can, ironically, be quite harmful to long-run growth*" (Stein, 1997. p. 284).

In the simpler part of his formal model, Stein (1997) assumes that only the entrants innovate and that there are no patents. In addition, there are no fixed costs and more importantly no R&D costs. Stein (1997) relaxes these assumptions separately in the latter part of his article. If only the entrant does research, the introduction of endogenous R&D spending intensifies the innovative waves. If the incumbent can also innovate, it is found that the probability of an entrant generating an innovation is still a decreasing function of the incumbent's age.

The two models that were summarized above have certain differences that are illustrative of the role that incremental innovation plays in the economy. Stein (1997) examined evolution of a single industry with two firms, an incumbent and a challenger, while Young (1993b) modeled ubiquitous quality improvements in the context of whole economies. Demand is largely absent from Stein's (1997) model. Yet, in Young's (1993b) model the size of a market, i.e. potential demand, creates a situation where the parameter concerning learning has no effect on the overall rate of growth. Furthermore, Stein's (1997) model has a firm and a product specific learning-by-doing, while Young's (1993b) model centers around product specific learning-by-doing. While these models can be illustrative of incremental innovation, neither of them is explicitly suitable for the examination of patenting.

Appendix C – Test of incrementality in a scale-intensive industry measured by the technological content of utility models

This appendix will this the incrementality of utility models. Because of the case specific hierarchical nature of incremental innovations, the incrementality of utility models must be tested in a certain kind of an industry and technology. The industry that is chosen for the purposes of this case study is NACE 28.95: Manufacture of machinery for paper and paperboard production. This industry manufactures *machines* for paper production. On the other hand, this industry could also be identified through IPC system – the code D21F indicates paper-making machines and methods of producing paper on them. Both routes of identification give the same results: 400 UM mentions and 1080 patent mentions of D21F.

The hierarchical nature of IPC was explicated in chapter 4.2. In the case of D21F, this subclass symbol denotes the highest level of the underlying group hierarchy: It is followed by groups which are denoted by the following kind of symbol, "number/00" (there are 8 of such classifications in existence), which in turn are dived into subgroups denoted by codes such as "a number/a number". These subgroups gestate their own hierarchies where some codes are ancillary to other symbols (generic meaning is denoted by a single dot in the exposition, multiple dots express *kinds* of the generic meaning).

As is already mentioned, right away it can be discerned that a larger proportion of patents is characterized simply by the subclass symbol without any group symbol. This alone could mean that patents involve higher levels of system hierarchy than utility models. However, differences in the sources of data and the fact that process inventions can only be patented with a patent hinder the correctness of such conclusions. The second observation is that utility models are associated with 41 and patents with 59 of the overall IPC groups in this industry, and the correlation between these two counts amounts to mere 0,2983. Again, the reason for this might lie in process inventions.

All IPCs (D21Fs) were lined up and were assigned numbers (1-4) indicating their position in the IPC hierarchy. In addition, IPCs were also assigned a process invention dummy which was the result of reading through IPC descriptions. The resulting correlations are summarized in table C.1. Bracketed correlations are calculated without the mere subclass description (D21F and nothing more). The technological content of utility models is much like the technological content of patents if mere sub-classes symbols are left out (correlation of 0.819) in paper-making machine industry. Both patent counts and utility model counts decrease as the hierarchical indicator decreases, however, rather erratically (correlations of -0.224 and -0.182).

Table C.1: Technological content of paper-making machines indicated by UMs and patents. (61 observations)

	UM	Patent	Hierarchy	Process dummy
UM	1			
Patent	0.298 (0.819)	1		
Hierarchy	-0.202 (-0.253)	-0.5 (-0.383)	1	
Process dummy	-0.217 (-0.224)	-0.139 (-0.182)	0.059 (0.038)	1

Table C.2 summarizes the regression results that were calculated with the same data set. Again, bracketed coefficients, numbers, and such indicate the regression that was conducted with the data without the mere subclass description. Utility model counts are explained with three regressors: patent counts, IPCs hierarchical position, and with a process dummy. The variable that is significant on 1% level depends on whether mere sub-class symbols are included in the data set: if they are, process dummy is the variable with the largest explanatory power, if they are not, patent counts have the largest explanatory power. Overall, the regression with mere sub-class symbols explains mere 0.1246 of the observed dispersion. However, if mere sub-class symbols are remitted the regression explains 0.6802 of the observed dispersion.

It is evident that utility models are not defined by their position in the system hierarchy indicated by IPC symbols. Instead, the overall inventive activity seems to explain the allocation of these petty patents, as well as, patents in the manufacture of machinery for paper and paperboard production industry (NACE 28.95). A large proportion of patents are assigned a more general IPC classification. Without this generic classification, utility models and patents involve similar inventions. Hence, it can not be stated that utility models indicate any more incremental inventions than patents do in this industry, that is, if the generic subclass description is remitted. Otherwise, utility models are best explained (badly) by their non-process-like inventive content. It should be remembered, however, that these results are likely to be driven by industry specific characteristics. Hence, any general conclusions about the utility model and its incrementality are of coarse imprecise.

Table C.2: Regression of UM characteristics
Dependent variable: UM count

Regressor	**Coefficient**	**Robust std. error**	**t**	**P>|t|**
Patent count	0.07 (0.51)	0.1 (0.11)	0.66 (4.61)	0.512 (0.000)**
Hierarchy no.	-0.99 (1.00)	1.68 (1.17)	-0.59 (0.86)	0.558 (0.395)
Process	-4.66 (-1.93)	1.38 (1.29)	-3.37 (-1.50)	0.001** (0.140)
Constant	9.42 (-3.05)	5.95 (4.35)	1.58 (-0.70)	0.119 (0.486)

Observations: 62. F(3, 58) = 4.43. Prob > F = 0.0072. R-squared = 0.1246. Root MSE = 9.889. [Observations: 61. F(3, 57) = 9.83. Prob > F = 0.0000. R-squared = 0.6802. Root MSE = 6.0096.]