2. Research Clustering, Co-Inventor Networks and Innovative Places: A Literature Survey

2.1. A Survey of the Theoretical Literature

2.1.1. The Co-Evolution of Research Strands in the Cluster Literature

As centuries of research have been dedicated to land-use, core-periphery structures and industry location, the research of economists and geographers has led to a broad and complex body of literature (Feldman, 1999; Harris, 2008; Thisse, 2011).[29] Therefore, these theoretical contributions need to be classified into several research lines, which will subsequently be presented and discussed in the following sections, even though the main focus of this study is on research clustering and the distribution of inventorship activity.

One suggestion regarding the origin of regional disparities and the causes and effects of clustering can be found in the work of Ottaviano and Thisse (2001, 159), who argued that

> "[i]f we want to understand something about the spatial distribution of economic activities and, in particular, the formation of major economic agglomerations [...] we must assume either (i) that *space is heterogeneous* (as in the neoclassical theory of international trade or in land-use models à la von Thünen), or (ii) that *production and consumption externalities exist* and are many (as in modern urban economics) or (iii) that *markets are imperfect* (as in spatial competition theory or in economic geography)."[30]

Today, it is quite common in the literature to separate the aforementioned elements into the following groups of the causes and effects that determine the (re-)location, co-location and co-agglomeration of research and production activities: (i) comparative advantage; (ii) externalities; and (iii) imperfect competition (i.e., monopolistic and oligopolistic competition) (Combes *et al.*, 2008, 42). However, the history of the literature on clustering, agglomeration economies and regional disparities can be traced back to the beginning of the 19th century. Nevertheless, for a long time, the processes of agglomeration and concentration occupied an inferior position on research agendas in the field of economics, and especially the role of spatial proximity and concentration with regard to innovative capacity, inventorship activity and innovation. Researchers have recently focused their attention on the conceptual relationship between agglomeration and concentration tendencies and the established spatial convergence studies that can be regarded as their empirical counterpart

---

29 See also Scott (2000), Clark *et al.* (2000) and Sheppard (2000).

30 The main advantage of the second source, namely technological externalities, is that this concept is compatible with perfect competition.

(Martin and Ottaviano, 2001; Brakman and van Marrewijk, 2008; Thisse, 2011).[31] Agglomeration and spatial concentration is nowadays increasingly challenged in economic theories (Rosenthal and Strange, 2001, 2003; Henderson, 2003a) and empirical analysis (Brülhart and Traeger, 2005; Geppert *et al.*, 2006; Fornahl and Brenner, 2009). Furthermore, the issues associated with divergence and spatial clustering have become highly visible with the emergence of metropolises, industrial belts and urban areas all around the globe, meaning that the world is considered to have become more "spiky" (Fujita and Krugman, 2003; Crescenzi and Rodríguez-Pose, 2008; Brakman and van Marrewijk, 2008).[32] As has been argued by many researchers, the earliest approaches and concepts date back to the seminal contributions made by Ricardo (1821), Launhardt (1882), Marshall (1920b), among others, and back to German location theory in particular (Thünen, 1966; Weber, 1929; Christaller, 1933; Lösch, 1954).[33] Regarding the different sources of agglomeration economies and factors that influence industry concentration and clustering of knowledge-intensive tasks, the work of Alfred Marshall (1890) is considered to be of central importance for both geographers and economists; especially the so-called "Marshallian externalities of the third kind" (Breschi and Lissoni, 2003; Press, 2006a; Capello, 2007).[34] A few decades after these externalities were proposed, in the 1950s and 1960s, the neoclassical literature on location was classified as the successor of the aforementioned classical contributions (Roos, 2002b; Press, 2006a).[35] The neoclassical approaches (a.k.a. regional science) improved the classical concepts of location, co-location and land-use; however, most contributions were unable to explain the processes of co-location, agglomeration and dispersion by means of different centripetal and centrifugal forces (Cruz and Teixeira, 2007; Blum, 2008). Moreover, heterogenous space and region-specific set-ups were not considered to be central elements. An exception to this rule are "comparative advantage models," which are based upon physical geography and heterogenous spatial structures (i.e., natural endowments) and are well-known in trade theory and cross-country trade studies (Combes *et al.*, 2008; Krugman, 2009; Thisse, 2011).[36]

The second half of the 20th century was determined by a meaningful expansion of the literature on clusters and by "new regionalism" (Storper, 1997, 2000; Scott and Storper, 2003), even though the contributions placed emphasis on different factors and relationships and originated from different schools of thought (Maggioni, 2002; Press, 2006a; Capello, 2007).[37] In light of these different theoretical advancements, which resulted in what Scott (2000) called "the great half-century" in economic geography, a broad range of concepts can be observed in retrospect.[38] Many of the concepts which were proposed could be classified

31 For an overview refer to Krugman (1992), Ellison and Glaeser (1997), Arbia (2001) and Baldwin and Martin (2004).

32 Refer also to Fujita and Mori (2005).

33 For a detailed literature survey, refer to Scott (2000), Keilbach (2000), Marques (2001), Martin and Sunley (2003), Jonas (2005), Press (2006a), Cruz and Teixeira (2007), Capello (2007), Eckey (2008), Freund (2008), Blum (2008) and Thisse (2011). For a comprehensive review and discussion of the growth and development theories between the 1950s and 1980s refer to Hagemann (2006) and Capello (2007).

34 For an overview refer to Keilbach (2000) and Cruz and Teixeira (2007).

35 See Isard (1956), Myrdal (1957), Alonso (1964) and Perroux (1970). For an overview refer to Keilbach (2000), Roos (2002b) and Press (2006a).

36 For an overview refer to Roos (2002b) and Capello (2007).

37 See also Scott (2000), Roos (2002b) and Martin and Sunley (2003).

38 Refer also to Scott (2000) for a detailed survey of the intellectual history of economic geography.

as knowledge-based cluster approaches, including the following: the well-known "Italian industrial districts" or "neo-Marshallian industrial districts" which consist of small Italian manufacturing firms with a region-specific tradition (Brusco, 1982; Becattini, 2002); the Californian School of geographers and their concept of "new industrial spaces" which focuses on the division of labor, vertical disintegration, transaction costs and path dependencies (Scott, 1988; Storper, 1997; Scott, 2000); the French approach of the "innovative milieus," developed by the GREMI group, which focuses on territorial specificities, networks and untraded interdependencies with respect to the innovation process (Aydalot, 1986; Camagni, 1991b, 1995; Capello and Faggian, 2005); the "neo-Marshallian nodes" (i.e., clusters) that are integrated into a global production network (Amin and Thrift, 1992); the "Porterian industrial cluster" approach, which is well-known from STI policy (Porter, 1998a; Delgado *et al.*, 2010); the "learning regions" approach that revolves around learning processes, networks and region-specific factors (Florida, 1995; Asheim and Isaksen, 2002; Asheim and Gertler, 2005); and finally the "systems of innovation literature," i.e., regional innovation systems, which center around learning processes, region-specific aspects and the institutional characteristics that influence the innovation process (Cooke, 2001; Doloreux and Parto, 2005; Cooke, 2008).[39]

The concepts and approaches listed above have contributed to our understanding of clustering, co-location, agglomeration and co-agglomeration in a meaningful way. Although inter-regional linkages are considered in these conceptions in some way, spatial proximity has remained the primal source of cluster dynamics, as it facilitates face-to-face interaction, networking and the exchange of tacit and codified knowledge between agents in knowledge-intensive industries. Furthermore, geographers and economists have found their own specific research agendas, although there is a considerable overlap in several fields with regard to theorizing and empirical analysis (Rodríguez-Pose, 2010).

In economics, the growing interest of researchers (and politicians) in a neoclassical tradition has let to the emergence of the so-called "new economic geography" (Krugman, 1991; Krugman and Venables, 1995a) (i.e., "geographical economics"), "urban economics" (Fujita and Thisse, 1996; Combes *et al.*, 2008) and the "endogenous growth theory" (Romer, 1986, 1990b). The community of researchers on geographical economics has contributed with new generations of models of co-location and agglomeration, which are mainly built upon pecuniary externalities that lead to an ever-contracting space (Krugman, 1991, 2009).[40] As a consequence, location, co-location and relocation are modeled as the outcome of an optimization process. The distribution of economic activity, i.e., industry location, and the emergence of regional disparities are solely dependent on centripetal and centrifugal forces that lead to either a symmetric distribution or a core-periphery structure. Consequently, the salient feature of the new economic geography is the formalization of cumulative circular causality, based upon factor mobility, vertical linkages or the accumulation of capital. In comparison to the "geography of innovation" literature (Feldman, 1999), urban economics similarly challenges knowledge diffusion and technological externalities with special focus

[39] The emergence of national, sectoral, technological and regional conceptualizations of the innovation system approach is the outcome of an intellectual debate on spatial heterogeneity, system boundaries and perspective/dimension of analysis (Cooke *et al.*, 1997; Fischer, 2001; Edquist, 2005; Doloreux and Parto, 2005; Cooke, 2007, 2008). See also Christ (2007) for an overview.

[40] For a comprehensive review see Martin (1999), Krugman (2000), Fujita and Krugman (2003), Fujita and Mori (2005), Combes *et al.* (2008), OECD (2009a), Rodríguez-Pose (2010), Thisse (2011).

on city structures and regional spillovers (Duranton and Puga, 2001; Storper and Venables, 2003; Henderson, 2010).[41] Thisse (2011, 7) noted that

> "urban economics now has strong links to theories of social networks and other forms of local interactions, the urban neighborhood being the place where many non-market relationships are developed."

The contributions to the endogenous growth theory are heavily built upon the ideas of Arrow (1962a,b) and the seminal work of Romer (1986) and Lucas (1988).[42]

In opposition to the optimization approach in geographical economics, researchers from the "evolutionary economics" and "economic geography proper" traditions have challenged the existence of clusters, agglomerations and varying economic development paths of regions differently (Martin, 1999; Garretsen and Martin, 2011).[43] They point to varying technological regimes, path dependencies, populations of heterogenous agents (e.g., entrepreneurs and researchers), institutional differences at the regional, sectoral and national levels, cultural differences and the informal institutions and networks that shape the regional landscape (Sheppard, 2000; Scott, 2000; Press, 2006a).

In an R&D clustering context, the knowledge-, learning- and innovation-based approaches to clustering, agglomeration and growth (a.k.a. "geography of innovation" literature) place the emphasis exclusively on the transmission and diffusion of channels of knowledge, interpersonal relationships in networks, interactive learning processes, cultural and institutional factors, and also on the sociological and behavioral aspects of the innovation process (Feldman, 2000; Bathelt *et al.*, 2004; Lundvall, 2007).[44] Moreover, one line of research focuses in particular on the distribution of creativity, i.e., researchers and entrepreneurs (Fujita and Thisse, 1996; Andersson *et al.*, 2005; Fingleton *et al.*, 2007). In addition, empirical studies nowadays focus on the mobility (and migration history) of researchers (Almeida and Kogut, 1999; Saxenian, 2006; Breschi and Lissoni, 2009), on co-patenting activities between individuals, firms and regions (Maggioni *et al.*, 2007; Maggioni and Uberti, 2009; Kroll, 2009), and on "paper trails" of knowledge diffusion that are measured by using patent citations (Jaffe *et al.*, 1993; Maurseth and Verspagen, 2002; Paci and Usai, 2009).

Recent analyses have placed a special emphasis on the formal and informal linkages between agents and different forms of proximity (spatial, technological, cultural). Consequently, the exploration of intra- and inter-regional networks lies at the heart of recent studies (Bergman, 2009; Burger *et al.*, 2009; Wilhelmsson, 2009).[45] As it is argued, the economic geography of innovation is increasingly combined with the concept of "open innovation," which encompasses the analysis of spatial knowledge domains, the outsourcing and fragmentation of R&D, and the transfer of different forms of knowledge across sectors and space (Cooke, 2007; Powell and Giannella, 2010).[46]

41 See also Glaeser (2000), Puga and Duranton (2000) and Combes *et al.* (2008).

42 For an overview refer to Grossman and Helpman (1991a), Grossman and Helpman (1994), Rima (2004), Jones (2004), Chandra and Sandilands (2005), Solow (2007), Harris (2008) and OECD (2009a).

43 See also Boschma and Frenken (2006) and Boschma and Frenken (2009a).

44 For further discussion and reviews refer to Feldman (1999), Asheim (2000), Audretsch and Feldman (2004), Jonas (2005), Press (2006a), Cruz and Teixeira (2007) and Blum (2008).

45 See also Porter *et al.* (2005), Capello (2007) and Bergman and Maier (2009).

46 Capello (2009) offered a comprehensive review of the conceptual approaches to agglomeration economies and develops a diachronic perspective; she argues that the 1970s were dominated by the

A complementary view that links several of the aforementioned approaches and epistemic communities is the "agglomerations and networks" line of research (Powell and Grodal, 2005; Johansson, 2005; Breschi *et al.*, 2005).[47] Today, geographers and economists are both increasingly emphasizing the idea that spatial interaction, networks and places are the key factors of regional development and key units of empirical analysis (Overman, 2003; Rodríguez-Pose, 2010; Garretsen and Martin, 2011). The central merit of these approaches is their specific emphasis on the identification and explanation of different effects and working channels of knowledge transmission, unintentional knowledge spillovers and network linkages in a spatial context.

In summary, for both economists and geographers alike, the world is considered to have become more "convex" and "spiky," and to be dominated by large and increasingly interconnected agglomerations separated by growing economic "deserts" (Florida, 2005; Duranton and Rodríguez-Pose, 2005; Rodríguez-Pose, 2010).

Unfortunately, a comprehensive review and discussion of all of the aforementioned theoretical concepts, approaches and research streams in the context of (research) clustering is clearly beyond the scope of this thesis.[48] Despite their general methodological and conceptual differences, meaningful overlaps and conceptual similarities between the aforementioned concepts and approaches can be observed, which will be illustrated in the subsequent theoretical review.[49]

In the following, the literature review summarizes major working channels and forces that determine clustering and agglomerative tendencies. A special emphasis is placed on research clustering and regional disparities of research and patenting activity. The empirical review in section 2.2 builds upon the theoretical survey and represents the starting point of the empirical analyses in this study.

2.1.2. From First-Nature Agglomerations to Knowledge-Intensive Industries

First-nature causes of co-location and agglomeration emerge from physical geography and are thus related to land use, climate, navigable waterways, immobile production factors

"industry" dimension, the mid-1970s by the "socio-cultural" dimension; the 1980s by the "cognitive" dimension, the 1990s solely by the "spatial" dimension, the late 1990s by the "geographic/industry" dimension and finally the 2000s onward by an "integrated approach" (Capello, 2009, 148).

47 See also Breschi and Lissoni (2009) and Bergman (2009).

48 The interested reader will find detailed reviews and surveys of the entire body of literature on clusters in, e.g., Feldman (1999), Roos (2002b), Maggioni (2002), Bathelt and Glückler (2003), Press (2006b) and Capello (2009), Rodríguez-Pose (2010), Garretsen and Martin (2011), Thisse (2011), among others.

49 For discussion and reviews refer to Overman (2003), Duranton and Puga (2004), Duranton and Rodríguez-Pose (2005), Polenske (2007), Capello (2007) and Harris (2008). Duranton (2008b, 10) has argued that *"[t]he relationship between [economic geographers and economists] has been fraught with difficulties. On the one hand, many geographers react very negatively to the renewed interest by economists in spatial issues. On the other hand, economists tend to ignore the work done by economic geographers. Despite these difficulties, geographers may learn something from the economists' more rigorous approach while the greater breadth of geographers may offer a great source of inspiration for economists."* For further comparisons of the different research communities see Castellacci (2007), Castellacci (2008), Rodríguez-Pose (2010).

such as labor and natural resources, among others. The spatial typology and regional (natural) endowments cannot be transformed or substituted (Mellinger *et al.*, 2000; Venables, 2006; Sachs and McCord, 2008).[50] Therefore, a region's natural endowments can be considered to significantly influence agglomeration and co-agglomeration of industries but also the spatial structure of innovative activity.[51] Accordingly, comparative advantage emerges from the heterogeneity of space and thus presupposes an uneven distribution of technologies, natural endowments, assets, or agents (Acs, 2002; Roos, 2002a; Combes *et al.*, 2008). Regarding physical geography, unfavorable characteristics of a location might be a mountainous surface and geographic remoteness, which provide inferior infrastructure potentialities, suboptimal climate conditions; favorable ones are, e.g., the availability of mineral resources, fertile soil, navigable seaways (rivers and harbors) (Eaton and Kortum, 2002; Roos, 2002b,a; Puga, 2010).[52] Trade structures are seen as the outcome of physical geography as has been reviewed by, e.g., Crafts and Venables (2003). As Hinloopen and van Marrewijk (2004, 3) stated,

> "the wood industry is usually located in areas with lots of trees; big harbors are usually at the mouth of a navigable river."[53]

In a US context, Ellison and Glaeser (1999) discussed the natural advantage (abundance) of the Washington state area in low-cost hydroelectric power, which has led to a significant co-agglomeration of energy intensive industries (see also Acs, 2002, 3). Accordingly, exogenous location determinants, which are not influenced endogenously by locations and their agent structure, represent a first group of exogenous agglomeration and location factors that can lead to core-periphery structures. The economic literature defines such factors (and effects) as "first-nature" agglomeration effects (or causes) as has already been stated by Marshall in his *Principles of Economics* ([1890] 1920). Marshall himself regarded such first-nature advantages as an important attribute for location in a historical context (see also Roos, 2002b, 66). In this respect, first-nature causes represent an origin of spatial heterogeneity.[54]Researchers have tried to measure the distribution of industries but also to quantify the importance of such exogenous location factors, especially related to issues of industrial production and the specialization of industries. Head *et al.* (1995) related the location decision of agents to spatial factor endowments, which represents the classical

[50] Gallagher (2008), among others, differentiated between first-degree linkages (transaction costs) and second-degree linkages (knowledge spillovers, labor pooling, input/market sharing, and natural advantage), which differs slightly from the applied classification in this study. For further ideas refer to Rosenthal and Strange (2004).

[51] The Heckscher-Ohlin framework and modeling alternatives of international trade are about first-nature causes of specialization in production (but not about agglomeration per se). On the basis of input endowments, these models are able to demonstrate why firms in one region tend to produce labor intensive, and in another region capital intensive goods. For a comprehensive review of the trade theory in economics see, e.g., Harris (2008) and Hofmann (2009).

[52] A comprehensive review of the relationship between geography and development and the importance of physical geography and continental patterns for location and co-location can be found in Mellinger *et al.* (2000).

[53] To give an additional examples: Napa Valley (California) has a specific climate, which is conducive to the harvesting of grapes and other fruits. Thus, the location is today a central node for the US wine and fruit industry, which co-agglomerate in the same location (Acs, 2002; Gallagher, 2008).

[54] As Marshall has emphasized *"[t]he chief causes [of industry localization] have been physical conditions; such as the character of the climate and the soil, the existence of mines and quarries in the neighborhood, or within easy access by land or water"* (Marshall, [1890] 1920, 269).

idea of endowment-driven industry location (see also Feldman, 2000). In the same line, Ellison and Glaeser (1999) showed that about one-fifth of spatial clustering of US-industrial production can be explained by an (even incomplete) set of natural advantage. Audretsch and Feldman (1996, 268) similarly reported evidence for centripetal forces originating from natural endowments that are independent from the industry life-cycle. They argued that

> "[t]he positive and statistically significant coefficients of natural resources suggest that a high dependence on natural resources tends to result in a greater geographic concentration of production in all four of the [industry] life cycle phases."

According to Acs (2002, 3), the former strength of the legendary US manufacturing belt in the northeastern and eastern part of Americas midwest (Wisconsin, St. Louis, Baltimore, Maine) was primarily based on physical geography, such as iron intermediates from Minnesota, coal inputs from mountains and water inputs from places nearby. Besides Ellison and Glaeser (1999), also Rosenthal and Strange (2001), Audretsch and Feldman (1996), and Kim (1995), among others, discussed natural endowment abundance as a significant driver of agglomeration and co-agglomeration. However, according to Ellison and Glaeser (1999), these factors can only explain 20-50% of industrial concentration. As a consequence, it can be argued that the concentration of industries is not solely determined by physical geography, i.e., first nature (Ottaviano and Thisse, 2004; Head and Mayer, 2004; Holmes and Stevens, 2004).[55] According to the issues raised by Acs (2002, 2), it is therefore essential to analyze in a European context (i) if industries and knowledge-intensive tasks are still highly localized in a few locations, (ii) if they relocate, to explain why they move to other locations or why activity shows dispersion, (iii) how firms and entire industries can (frequently) relocate with parts of their knowledge base. Obviously, additional drivers of agglomeration and dispersion seem to exist. Similar to Acs (2002), Krugman (1992, 5) was asking why the spatially concentrated US manufacturing belt could persist for such a long time, although the gravity centers of mineral products and other inputs have relocated.[56] The main objective of this thesis is to analyze the distribution of patenting activity in Europe and to explore the spatial structure of co-patenting activity.[57]

With regard to technological progress, Sachs and McCord (2008) argued that advancements in telecommunication technology are significantly affecting the global division of labor and the nature of agglomeration economies, which should give rise to "secondary growth poles." Regarding patenting and research activity, such developments might also exist in a European context, which will be addressed in chapters 3 and 4. Nevertheless, location still seems to matter a lot. In the US context, it has been argued by Acs (2002), among others, that many processes within the manufacturing value chain can still exist in the neighborhood because of an established regional knowledge base that has induced agglomerative forces by itself (see also Krugman, 1995; Audretsch, 1998; Klein and Crafts,

55 See also Combes and Overman (2004).

56 Krugman (1992, 5) argued: *"Think of the United States: most of the population of huge, fertile country lives along parts of two coasts and the great lakes [...]. [T]hese urban areas in turn are highly specialized, so that production in many industries is remarkably concentrated in space. This geographic concentration of production is clear evidence of the pervasive influence of some kind of increasing returns."*

57 For a detailed discussion refer to Glaeser *et al.* (1992), Jaffe *et al.* (1993), Rosenthal and Strange (2001), Cappellin (2001), Roos (2002b), Johansson and Quigley (2003), Abreu *et al.* (2004), Henderson and Thisse (2004), Ottaviano and Thisse (2004) and Klein and Crafts (2010).

2010). However, the relocation of some gravity centers has also initiated the creation of new high-technology industries and modern knowledge bases in other areas. One of these high-technology locations became known as the popular Silicon Valley in California. According to Acs (2002), most registered US inventors (and applicants of patent applications) in high-technology are today located in a few famous high-technology agglomerations and centers of research excellence, such as Cambridge, Massachusetts and Silicon Valley, but not in the former industrial centers, e.g., Detroit, Cleveland, Dayton (see also Audretsch and Feldman, 1996).[58] With regard to the aforementioned aspects, Florida (2002b, xi) noted that

> "[t]he new geography [...] is not the result of natural endowments of land, labor and capital [...]. Rather, [...] it is powered by innovation and entrepreneurship; and this in turn is the product of real people acting in real places. In other words, the factors that really matter are the ones we create for ourselves. That is because they are able to attract, mobilize and connect the factors that really matter - innovative people and creative entrepreneurs. [...] It was clear to me and to others that innovation is a geographically concentrated process; and there were certainly studies of this. But no one had really nailed it down. A big piece of the problem was that the field lacked the kind of measures required to probe this issue."

In the context of knowledge-intensive industries, one promising line of reasoning and currently popular line of research grounds on the assumption that knowledge bases are becoming increasingly global and mobile, which implies that research activities and co-inventor network linkages take place at a distance and that inventor networks frequently relocate in space (Breschi and Lissoni, 2009).[59] Therefore, this development is regarded as a fundamental change in the geography of innovation because research collaboration linkages between agents and firms are becoming increasingly border-crossing and international (Maggioni and Uberti, 2009; Hoekman *et al.*, 2010; Powell and Giannella, 2010).[60] Such developments should change the distribution of patenting activity and the spatial structure of co-patenting linkages. The thesis challenges this idea empirically in a pan-European context.

2.1.3. Agglomeration, Indivisibilities and Fragmentation

Modern literature on agglomeration economies is considered to represent a collection of reinterpretations and formalizations of different dimensions on the micro foundations of

---

[58] Acs (2002, 4), among others, reported empirical evidence for this hypothesis by contrasting inventorship activity of the American sunbelt states and the former industrial heartland. Today, the leading innovative US regions are Santa Clara (CA), Los Angeles (CA), Cook (IL), Middlesex (MA), Norfolk (MA), Orange (CA), Bergen (NJ), New York (NY), Fairfield (CN), Nassau (NY), Dallas (TX), San Diego (CA). It is clearly visible that not only most innovations come from states such as California, Massachusetts, New York and New Jersey, but also that these US states provide the majority of population employed in high-tech manufacturing and knowledge intensive services (see also Audretsch and Feldman, 1996, 2004).

[59] For an overview refer to Almeida and Kogut (1999), Agrawal *et al.* (2006) and Oettl and Agrawal (2008).

[60] It should be possible to recognize the spatial shift of inventorship activity in patent documents, i.e., the relocation of innovative activity in terms of inventorship relocation.

increasing returns and agglomeration economies from the last century (Duranton and Puga, 2004; Combes and Overman, 2004; Capello, 2009).[61]

According to Starrett's "spatial impossibility theorem," any competitive equilibrium will feature autarchic locations under the assumption of homogeneity of space, without increasing returns or indivisibilities, and the presence of transportation costs (Starrett, 1978).[62] Starrett (1978, 27) argued that,

> "[a]s long as there are some indivisibilities in the system (so that individual operations must take up space) then a sufficiently complicated set of interrelated activities will generate transport costs."

Fujita and Krugman (2003), among others, pointed out that the competitive framework can, however, not explain the occurrence of agglomerations in a closed, homogeneous space under constant returns to scale (CRS) production technologies without first-nature heterogeneity, as described in the last section, and/or indivisibilities.[63] Otherwise, increasing land rents would lead to a dispersion of production activity. However, as soon as economic activities are not perfectly divisible, they have a certain (sustainable) location (Fujita and Krugman, 2003; Duranton and Puga, 2004; Behrens and Thisse, 2006).

Capello (2009) argued that the concept of agglomeration economies can generally be classified into three micro-foundations: (i) indivisibilities, (ii) synergies and (iii) spatial proximity (see also Capello, 2007). The concept of indivisibilities is generally built upon an industrial (only implicitly geographic) dimension and places emphasis on productivity effects that originate from large-scale production and shift a firm's production or cost curve (Edwards and Starr, 1987; Rosenthal and Strange, 2001; Roos, 2002b).[64]

In an urban economics context, an evident cause of concentration of production and co-location of different firms is based on the advantages associated with the division of labor, which allows specialization and large-scale production yielding lower costs per unit of output (Duranton and Puga, 2004; Combes *et al.*, 2008; Capello, 2009). Duranton and Jayet (2005) differentiated between two strands of the literature on the division of labor. In a continuous framework, labor specialization can become ever narrower as the market size increases and the fragmentation of tasks increases with the market size. If, however, the

61 See also Rima (2004), Chandra and Sandilands (2005) and Combes *et al.* (2008).

62 For an overview see Ottaviano and Thisse (2001), Duranton and Puga (2004), Combes *et al.* (2008), Baumol (2008) and Puga (2010). Refer also to Ottaviano and Thisse (2000), Fujita and Krugman (2003) and Duranton (2008b).

63 It is generally assumed that goods are consumable in infinitely divisible quantities. On the production side of the market, indivisible equipment is identical to "fixed costs" and inputs only exist in minimum quantities. Thus, indivisible inputs are associated with scale economies on the production side. Moreover, if the indivisible inputs are not overly specialized, they can be implemented in diversified processes at a lower cost compared to separately specialized plants, i.e., economies of scope (Edwards and Starr, 1987; Duranton and Puga, 2004; Combes *et al.*, 2008; Duranton, 2008b).

64 Similarly, Kaldor (1966) has addressed this idea by distinguishing between increasing returns in the manufacturing industry and decreasing returns in the primary sector what became known under the label Verdoorn's law or Kaldor's second growth law (see also Seiter, 1997; Rima, 2004; Seiter, 2005; Capello, 2007; Combes *et al.*, 2008; Capello, 2009). Kaldor (1972, 1243) mentioned the importance of Allyn Young's contribution to the debate: *"[Young's article] was so many years ahead of its time that the progress of economic thought has passed it by [...] partly because its criticism of general equilibrium theory could not be appreciated at the time when that theory itself was not properly understood."*

benefits from the division of labor originate from small indivisibilities at the worker level, then the fragmentation of tasks proceeds discontinuously as the market size increases. Employing a worker in different tasks would cause set-up costs as it prevents specialization. Given a sufficient scale, it is preferable to allow the fragmentation of tasks and thus labor specialization, which avoids switching costs. In a spatial context, the presence of more workers in a given activity within a location may increase the output more than proportionately as it allows them to specialize in a narrower set of tasks (Duranton and Puga, 2004; Kim, 2006; Puga, 2010).[65] Consequently, if the division of labor (i.e., vertical disintegration and fragmentation of tasks) is significantly distance-sensitive and/or shows other forms of indivisibilities, then it becomes clear why co-location and co-agglomeration are beneficial for agents.[66]

In a similar way, Combes *et al.* (2008, 39) argued that the existence of non-ubiquitous agents (i.e., human capital) and scale economies can be interpreted as specific forms of indivisibilities. However, an increasing sub-division/fragmentation of tasks (i.e., division of labor) raises scale economies and the heterogeneity of skilled labor, which may also increase labor-matching costs.

Duranton and Puga (2004, 2065) brought forward the critique that it is hard to think of any single activity or facility subject to meaningful indivisibilities to justify the emergence or existence of cities and metropolises. Accordingly, they present three mechanisms in the context of agglomeration economies. Spatial proximity helps in (i) sharing, (ii) matching and (iii) learning.[67] A larger market allows sharing mechanisms that cover the sharing of indivisible facilities, the sharing of gains from a wider variety of input suppliers in a location, sustained by a larger final good industry, the sharing of gains from individual specialization, sustained by larger production, the sharing of a local labor market and finally the sharing of risk. Matching mechanisms, on the other hand, are the improvements of the probability and quality of matching between agents that originate from large (dense) markets, e.g., employees and employers (i.e., labor market externalities), partners in joint projects, or financiers and entrepreneurs. Generally, matching mechanisms can be related to workers, intermediates and ideas. Finally, learning mechanisms are related to the generation, diffusion and accumulation of knowledge in a spatial context (i.e., innovation externalities), e.g., learning about market evolutions, new technologies and new forms of organization and routines (Duranton and Puga, 2004; Duranton, 2008a; Capello, 2009). Similarly, Florida (1995, 531) has argued that

> "[t]he shift to knowledge-intensive capitalism goes beyond the particular business and management strategies of individual firms. It involves the development of the new inputs and a broader infrastructure at the regional level on which individual

65 The gains of the division of labor are well-known since Adam Smith (Roos, 2002b; Capello, 2009). The idea of specialization by division of labor in a spatial context has also been mentioned by List ([1842] 1909). A few decades later, Young (1928) extended the discussion about increasing returns by addressing Smith's division of labor concept (see also Rima, 2004; Seiter, 2005; Chandra and Sandilands, 2005).

66 Refer to Edwards and Starr (1987), Duranton and Puga (2004), Press (2006a) and Duranton (2008b) for a review and detailed discussion of indivisibilities, specialization and the division of labor. The whole debate on increasing returns is beyond the scope of this study.

67 Refer also to Capello (2007, 2009) and World Bank (2009) for an overview.

> firms and production complexes of firms can draw. The nature of this economic transformation makes the regions key economic units in the global economy."

To conclude, the firm's incentive to concentrate all its production in a single location is, however, not identical to the advantages that originate from proximity to other firms and the advantages of fragmentation in local markets. It is argued that large-scale production promotes returns internal to the firm, which represent a first meaningful incentive for a firm to concentrate production in a single location. Second, firms are considered to (re-) locate production close to a large market in the case of significant costs of transportation. Third, co-location and co-agglomeration in cities and large urban areas are preferred because fragmentation of production induces essential input-output linkages that affect input prices (Gallagher, 2008; Combes *et al.*, 2008; Capello, 2007, 2009).[68] Spatial proximity is assumed to induce additional external economies which are independent from internal scale economies but originate from the scale of the local market (Ciccone, 2008; Henderson, 2010). Therefore, central classifications of external economies will be reviewed and discussed in the subsequent section, particularly those related to knowledge-intensive industries and innovation.[69]

2.1.4. Agglomeration, Clustering and External Economies

2.1.4.1. Industrial Districts and External Economies

The external advantages of agglomerated activities and environments are considered to represent central determinants of the spatial distribution of production and research activities across regions. It is argued that regional disparities are persistent phenomena because of "second-nature" agglomeration economies (Fujita and Thisse, 1996; Acs, 2002; Duranton, 2008a).[70]

Long ago, Alfred Marshall ([1890] 1920b) has disclosed the advantages of co-location and spatial proximity. In the following, the main arguments are briefly reviewed. Marshall (1920b, 271) argued that

> "[w]hen an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from near neighbourhood to one another. The mysteries of trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously. [...] Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined

68 See also Krugman and Venables (1996), Holmes (1999) and Harris (2008). In an US context, Glaeser (2005a) suggested that economies in transportation may also explain why industries became concentrated in cities.

69 The following section places emphasis on Marshall's external economies put forward in his *Principles of Economics* ([1890]/1920), Chapter X of Book IV. Several economists equalize Marshall's arguments with a general Marshallian agglomeration theory (Roos, 2002b, 2008; Capello, 2007; Combes *et al.*, 2008).

70 See also Caniëls (1996), Keilbach (2000), Roos (2002b), Press (2006a), Capello (2007) and Fingleton (2007).

> with suggestions of their own: and thus it becomes the source of further new ideas. [...] and presently subsidiary trades grow up in the neighbourhood, supplying [the industry] with implements and materials, organizing its traffic, and in many ways conducing to the economy of its material."

Besides these input externalities, another agglomerative effect, which was addressed by Marshall, is the observed tendency of firms and entrepreneurs to locate near specialized markets for labor, what is nowadays discussed under the label "labor-market pooling" or "labor-market externalities" (Krugman, 1995; Combes and Duranton, 2006; Martin *et al.*, 2008).[71] As Marshall (1920b, 270) argued,

> "[a]gain, in all but the earliest stages of economic development a localized industry gains a great advantage from the fact that it offers a constant market for skills. Employers are apt to resort to any place where they are likely to find a good choice of workers with special skill which they require; while men seeking employment naturally go to places where there are many employers who need such skill as theirs and where therefore it is likely to find a good market."

Marshall additionally addressed the consumption behavior of agents in a spatial context, i.e., local markets and local demand by consumers. This is similar to the new economic geography framework (section 2.1.5.5), where centripetal forces increase with the size of the local market (i.e., pecuniary externality). In the new economic geography, the process of agglomeration is enforced by cumulative causation and circular causality as production factors (and demand) are inter-regionally mobile (Krugman, 1991; Roos, 2002b; Capello, 2007).[72] Additionally, Marshall pointed to the effects of co-location on the transaction and search costs related to the consumer's preferences. In this respect, he implicitly discussed the benefits of co-location of specialized suppliers and vertical disintegration, although his remarks are related to consumers and not directly to intermediate industries.[73]

The aforementioned factors are considered to affect different levels of aggregation; i.e., the firm-/ plant-level level, the regional level, the industry-level. Some effects are surely external to single firms but internal to local industries, whereas Marshall's attention was primarily on "proximity externalities" in industrial district. Other effects are, however, external to the industry but internal to the (regional) economy as a whole. This makes several agglomeration effects (i.e., technological externalities) compatible with the perfect competition framework (see section 2.1.4.3).

As is frequently argued, literature on clustering and agglomeration uses Marshall's external economies as a main reference with respect to economies of localization and urbanization and the dynamic effects from agglomeration (Audretsch and Feldman, 1999; Capello, 2007; de Groot *et al.*, 2009).[74]

---

71 Refer also to Fujita and Thisse (1996), Roos (2002b), Sonobe and Otsuka (2006), Press (2006a), Capello (2007), Combes *et al.* (2008), Harris (2008), World Bank (2009) and Overman and Puga (2010) for an overview.

72 For additional overviews refer to Keilbach (2000) and Press (2006a).

73 *"[T]he consumer will go to the nearest shop for a trifling purchase; but for an important purchase he will take the trouble of visiting any part of the town where he knows that there are specially good shops for this purpose"* (Marshall, 1920b, 273).

74 Krugman used Marshall's agglomeration economies to overcome perfect competition. He modeled economies of scale internal to the individual firm (plant-level), which is based on the idea of imperfect

Related to the technological externalities debate, a central part in Marshall is devoted to the diffusion of ideas and knowledge in a spatial context, nowadays labeled "Marshallian externalities of the third kind" (Breschi and Lissoni, 2001a,b; Rosenthal and Strange, 2004). In this context, the diffusion of economically useful knowledge is the source of technological externalities, as the

> "[n]ew idea, [...] taken up by others and combined with suggestions of their own [...] becomes the source of further new ideas" (Marshall, 1920b, 270).

In *Industry and Trade*, Marshall (1920a, 190) has additionally argued that

> "[t]he leadership in a special industry, which a district derives from an industrial atmosphere, such as that of Sheffield or Solingen, has shown more vitality than might have seemed probable in view of the incessant changes of technique. The explanation is perhaps to be found in the fact that an established centre of specialized skill, unless dominated by a gild or trade-union of an exceptionally obstructive character, is generally in a position to turn to account quickly any new departure affecting its work; and if the change comes gradually, there is no particular time at which strong incitement is offered to open up the industry elsewhere."

Related to the previously mentioned aspects, Marshall attributed central importance to the "industrial atmosphere" in districts, which originates from the presence of skilled people that transform regions (and districts) into leading industrial places. In a more socio-cultural perspective, the Marshallian industrial atmosphere is also interpreted as the advantages that arise from (localized) networks and social proximity in urban areas (Fujita and Thisse, 1996; Capello, 2007, 189). The capacity of agents to co-operate is rooted in the socio-cultural environment which generates increasing returns, the so-called "district economies" (Roos, 2002b; Capello, 2007, 2009).[75] These economies are based on trust, sense and social proximity, which represent indivisibilities, and spatial proximity is considered a meaningful prerequisite (Capello, 2007, 2009).[76]

To summarize, it is argued that Marshall early contributed with a well-defined classification of external advantages of agglomerated activities and environments (Scitovsky, 1954; Fujita

competition. This is also a reason why non-pecuniary effects are not modeled in early new economic geography models as the novelty is related to pecuniary effects from increasing returns. Equilibrium city size (or more general the manufacturing share in the region) depends on the trade-off between pecuniary externalities (centripetal forces) and the costs of spatial concentration (centrifugal forces) (Feldman, 2000; Keilbach, 2000; Press, 2006a; Capello, 2007; Audretsch *et al.*, 2008; Combes *et al.*, 2008). As Krugman has argued: *"Thus local external economies never disappeared as a concept from economics. Indeed, if you were ask a mainstream economist at any time between, say, 1930 and the last few years why cities exist, or why some industries are so concentrated in space, he or she would surely answer in terms of just such local externalities [technological and pecuniary externalities based upon Marshall]"* (Krugman, 1995, 50).

75 Further to this, it is argued that Marshall linked geographical proximity with the transfer mechanisms of knowledge, which improves the level of productivity in all companies (Keilbach, 2000; Press, 2006a; Capello, 2007).

76 As Marshall (1920a, 189) has argued: *"[...] personal contact is most needed (1) in trade between allied branches of production, at all events in regard to things which have not yet been brought completely under the dominion of either general or particular standardization; and (2) in all dealings, especially retail, connected with dress, ornaments and other goods, which need to be adapted to individual requirements and idiosyncrasies."*

and Thisse, 1996; Rosenthal and Strange, 2004).[77] As a consequence, there exist several interpretations of Marshall's external economies, i.e., building blocks of agglomeration economies in the literature, that place emphasis on different working channels (Duranton and Puga, 2004; Capello, 2009).

2.1.4.2. Interpretations of Marshall's Agglomeration Economies

A popular interpretation of Marshall's external economies is given by Krugman (1991) and has been used in his new economic geography framework and in further framework advancements (see sections 2.1.5.5 and 2.1.6.7). According to Krugman, agglomeration economies can generally be classified into three forces: (i) human capital externalities, (ii) technological externalities and (iii) market interaction (pecuniary) externalities (Krugman, 1991, 1995; Capello, 2007, 2009).[78] Similarly, Martin *et al.* (2008) classified Marshall's agglomeration economies into (i) labor market externalities, (ii) knowledge externalities and (iii) input externalities. The latter force is also known as pecuniary externality. The incorporation of technological externalities was a general approach in endogenous growth models (see section 2.1.6.6), whereas pecuniary externalities were central features of the new economic geography. The application of both externality types to innovation clusters and industry agglomerations has recently (re-)accelerated. However, "technological spillovers," although recognized, are not included in Krugman's seminal core-periphery framework. In this respect, an illustrative statement on the importance of such spatial external effects and emphases in economic models was offered by Fujita *et al.* (2001, 4), although technological externalities are downplayed:[79]

> "[A]lthough all three of Marshall's forces are clearly operating in the real world, the new economic geography models have generally downplayed the first two [knowledge spillovers and markets for skilled labor], essentially because they remain hard to model in any explicit way. Instead they have focused on the role of linkages [backward and forward linkages associated with large local markets]."

---

77 See also Press (2006a), Martin *et al.* (2008), Combes *et al.* (2008), Capello (2007, 2009) and Puga (2010). Fujita and Thisse (1996, 345) concluded that, *"[f]ollowing Scitovsky (1954), it has been customary to consider two categories: 'technological externalities' (such as spillovers) and 'pecuniary externalities'. The former deals with the effects of nonmarket interactions which are realized through processes directly affecting the utility of an individual or the production function of a firm. By contrast, the latter refers to the benefits of economic interactions which take place through usual market mechanisms via the mediation of prices. For obvious reasons Marshall was not aware of this distinction, and his externalities turn out to be a mixture of technological and pecuniary externalities. As a consequence, each type of externality may lead to the agglomeration of economic activities."*

78 Krugman (1992, 36) argued: "*First, [...] a pooled market for workers with specialized skills; [...] Second, an industrial center allows provision of nontraded inputs specific to an industry in greater variety and at lower cost [...] Finally, because information flows locally more easily than over great distance, an industrial center generates what we would call technological spillovers.*" See also Crafts and Venables (2003) for an identical classification and review.

79 Another interpretation of Marshall's described economies, very similar to the one of Krugman, is given by Fujita and Thisse (2003, 8), who put emphasis on (1) internal economies, (2) the formation of a highly specialized labor force and the production of new ideas, both based on the accumulation of human capital and face-to-face communication, (3) the availability of specialized input services and (4) the existence of modern infrastructures.

Labor-market externalities, as described by Marshall, represent another salient feature of cities and agglomerations (Crafts and Venables, 2003; Combes and Duranton, 2006; Henderson, 2010). Rosenthal and Strange (2004, 2153) classified the gains and incentives of labor-pooling into two distinct aspects. First, when firms (industries) show an ex ante lack of knowledge about the labor market structure in a region, firms tend to agglomerate in urban areas and cities. Such locations are said to provide dense and heterogenous labor markets, which increase the probability of a better match of specific labor demand and supply. Therefore, growing cities show an increasing diversity of labor supply, which should lead to fulfilling the needs of knowledge-intensive industries. Second, firms with ex ante knowledge of regions' labor markets tend to co-agglomerate (locate) in cities or urban areas, which minimizes search costs, training costs and risks for employees. Furthermore, if industries are hidden by positive demand shocks, firms will generally find it easier to hire additional workers in urban regions and cities (Duranton and Puga, 2004; Gallagher, 2008; Martin *et al.*, 2008).[80] In a research clustering perspective, one may argue that regional disparities in research activity are (increasingly) dependent on the spatial distribution of human capital (i.e., researchers and creative minded people) which grants migration and network studies a pivotal attention (Florida, 1995; Fujita and Thisse, 1996; Breschi and Lissoni, 2009).[81]

In a more general way, Fujita and Thisse (1996) differentiated between market-led mechanisms and the effects that occur outside the anonymous market.[82] Fujita and Thisse (1996, 345) concluded that

> "[i]t is now customary to [distinguish between] two categories: technological externalities and pecuniary externalities. The former deals with the effects of non-market interactions that are realized through processes directly affecting the utility of an individual or the production function of a firm. In contrast, pecuniary externalities are by-products of market interactions [transactions]: They affect firms or consumers and workers only insofar as they are involved in exchanges mediated by the price mechanism. Pecuniary externalities are relevant when the markets are imperfectly competitive, for when an agent's decision affects prices, it also affects the well-being of others."

In contrast to Krugman (1991), Ottaviano and Thisse (2001) gave much more attention to technological externalities that originate from knowledge transmission at a proximate distance. However, they also placed a special emphasis on the spatial scale of analysis when differentiating between the possible externalities.[83]

In light of the aforementioned determinants and drivers of clustering and agglomeration, the role of knowledge externalities was (and is) widely discussed in literature, especially in the context of knowledge generation and diffusion in cities (Glaeser *et al.*, 1992; Fujita and Thisse, 1996; Johansson, 2005).[84] As this study is explicitly focusing on the determinants

80 See also Keilbach (2000), Roos (2002b), Capello (2007) and Combes *et al.* (2008).
81 See also Zucker *et al.* (1998) and Almeida and Kogut (1999).
82 For a further discussion refer to Ottaviano and Thisse (2001).
83 Refer also to Lucas and Rossi-Hansberg (2002), Duranton and Puga (2004), Boschma and Frenken (2009a).
84 Refer also to Henderson *et al.* (1995), Black and Henderson (1999b), Glaeser (2000), Duranton and Puga (2001), Henderson (2003a), Duranton and Puga (2004) for an overview.

of research clustering and the distribution of patenting activity, the access to information and knowledge and its transmission are assumed to represent pivotal factors of a firm's location decision. Related to this perception, Fujita and Thisse (1996) pointed to the meaningful differentiation between production and creation as these are distinct activities of individuals, whereas the existence of pecuniary externalities in agglomerations is a crucial factor for production processes in the manufacturing industry (and industrial geography). Creative activities of individuals are in particular influenced by their proximity to other people. As a consequence, economic activities in the "knowledge economy" are considered to particularly depend on creativity which is identical to Florida's "creative class concept" (Florida, 1995, 2002c,a; Florida and Tinagli, 2004).[85] Similarly, Lucas (1988, 39) argued

> "[t]hat the 'force' we need to postulate account for the central role of cities in economic life is of exactly the same character as the 'external human capital' I have postulated. [...] What can people be paying Manhattan or downtown Chicago rents for, if not for being near other people?"

Hence, personal communication and knowledge transmission within and between groups of individuals (and epistemic communities) are considered to be vital preconditions for creativity and innovation output in knowledge-intensive industries.[86]

Finally, in a core-periphery perspective, Kilkenny (2010) brought forward the argument that remote low-density areas (rural regions) seem to be competitively disadvantaged due to a significant lack of static and dynamic agglomeration externalities, which gives inter-regional research linkages a pivotal role in regional development.[87]

To sum up, knowledge externalities seem to matter for the development of new routines and new products. Neighboring regions are considered to benefit from spatial proximity to high-level growing regions if there are considerable positive externalities and inter-regional flows of knowledge. Moreover, firms and agents have access to knowledge bases via intra- and inter-regional research linkages and networks which also induce some kind of externalities. The following sections are organized in order to offer a detailed classification and taxonomy of the causes and effects of agglomeration, co-agglomeration and clustering. Special attention is drawn to research clustering, knowledge transmission and the spatial concentration of research and innovative activity.

2.1.4.3. Agglomeration Economies, Spillovers and Networks: A Taxonomy

It is evident from the previous sections that the origins of agglomeration economies and the causes of clustering are indeed multifaceted. The literature generally distinguishes between(i) intra-market externalities (pecuniary externalities) that work via prices, (ii) quasi-market externalities (externalities from a network transactions) and (iii) extra-market externalities (technological externalities) that occur without any monetary compensation

85 Refer also to Jacobs (1969) and Glaeser (2005a).

86 Similarly, Robert-Nicoud (2004, 4) argued that *"[a]gglomeration generates inertia [...], people and firms are there because other people and firms are there too. So people are willing to move out of the agglomeration [relocate] only if a large shrunk of people are willing to do as well."*

87 Refer also to Partridge and Rickman (1999), Duranton and Puga (2001), Duranton (2008a), de Groot *et al.* (2009).

(Johansson, 2005; Capello, 2007, 2009). Moreover, a general taxonomy can be built upon the following pillars: (i) the source of externalities (proximity vs. network link externality); (ii) the effects and consequences of externalities (efficiency vs. innovation externality); and (iii) the nature of externalities (pecuniary vs. non-pecuniary externality). The industry dimension on agglomeration economies is additionally considered with the concepts of urbanization and localization economies (section 2.1.5). Furthermore, the concept of innovation externalities (section 2.1.6), i.e., the concepts of Marshall-Arrow-Romer externalities (section 2.1.6.2), Jacobs externalities (section 2.1.6.3) and Porter externalities (section 2.1.6.4) are considered. Finally, special attention is given to the generation and transmission of knowledge via anonymous market transactions, via (persistent) inter-regional network linkages and intentional and unintentional knowledge flows in localized networks and industry clusters (section 2.1.7).

The subsequent table 2.1 illustrates a general typology of the aforementioned externalities (Johansson, 2005; Capello, 2007, 2009); it classifies horizontal and vertical externalities against efficiency and innovation externalities.[88] The taxonomy seems to fit to the research agenda of several epistemic communities; e.g., geographical economics, economic geography, evolutionary economics, evolutionary economic geography and geography of innovation. The different externality concepts are reflected in different models, which will be presented and discussed in the following sections.[89]

2.1.5. Agglomeration, Research Clustering and Pecuniary Externalities

2.1.5.1. Pecuniary Externalities, Local Scale and Efficiency

Pecuniary externalities are of broad interest in explaining the spatial concentration, agglomeration and co-agglomeration, as shown by the majority of new economic geography (NEG) models and approaches in economic geography proper.[90] They are also sometimes labeled vertical spillovers, welfare spillovers or rent spillovers (Johansson, 2005; Harris, 2008).[91]

[88] The classification also takes into consideration the following contributions: Scitovsky (1954), Puga and Duranton (2000), Duranton and Puga (2004), Jacobs (1969), Acs *et al.* (1997), Acs *et al.* (2002), Ottaviano and Thisse (2001), Kelly and Hageman (1999), Martin and Sunley (2003), Caniëls (2000), Glaeser *et al.* (1992), Glaeser and Resseger (2009), Henderson *et al.* (1995), Audretsch and Feldman (1995), Audretsch and Feldman (1999), Audretsch and Feldman (2004), Audretsch and Keilbach (2008), Glaeser (2000), Feldman (2000), Roos (2002b), Döring (2004), Autant-Bernard and Massard (2007), Autant-Bernard *et al.* (2007), Döring and Schnellenbach (2006), Athreye and Werker (2004), Breschi and Lissoni (2001b), Breschi *et al.* (2005), Press (2006a), Keilbach (2000), Maggioni (2002), de Groot *et al.* (2009), Andersson *et al.* (2005), Greunz (2005), Rosenthal and Strange (2001), Rosenthal and Strange (2003), Harris (2008).

[89] The following summary of models and concepts is, however, non-exhaustive.

[90] For an overview refer to Ottaviano and Puga (1998), Neary (2001), Duranton and Puga (2004), Autant-Bernard and Massard (2007) and Capello (2007).

[91] Verspagen (1997, 230) has argued that *"Griliches (1979) termed this form of spillovers 'rent spillovers', because they are crucially related to the rents of both the receiving and supplying firm. On a different, more semantic level, Griliches (1992) has argued that as long as goods are being traded between the supplying and receiving party, there are no 'real' externalities, in the strict sense of the word, involved. Although one might therefore argue that the term 'spillover' is less appropriate in this case, there is no need to abandon the terminology as long as it is clear that rent spillovers*

Table 2.1. Innovation vs. efficiency externalities

Form/type	Innovation externality	Pecuniary/efficiency externality
Proximity externality	Absolute size and/or diversity of local market affects product development (early phases of a product cycle) and knowledge spillovers on innovative output (MAR vs. Jacobs)	Size of local market induces scale economies for producers (distance-sensitive production)
Vertical	Downstream externality from knowledge flows between supplier and customer; proximity externality and/or network transaction externality	Downstream externality affecting the price (supplier, customer)
Vertical	Upstream externality from knowledge flows between input (knowledge) buyer and seller; proximity externality and/or network transaction externality	Upstream externality affecting input costs (of a company)
Horizontal	Knowledge flows between competitors from joint R&D efforts based on a transaction linkage (network linkage) or based on (unintended) spillovers in an agglomeration due to proximity	Co-operation between competitors (transportation, marketing, long-distance export)

Source: illustration based on Johansson (2005, 112); see also Johansson and Quigley (2003), Capello (2007, 2009) and Burger *et al.* (2009).

Pecuniary externalities are based on market interactions and affect firms (or consumers) by means of exchanges involving prices (Ottaviano and Thisse, 2001; Duranton and Puga, 2004; Johansson, 2005).[92] The interdependence between the supply side (firms, products) and demand side (consumers, market size) is direct due to the spatial range and location of pecuniary externalities (Autant-Bernard and Massard, 2007; Dewhurst and McCann, 2007). NEG models, for example, treat market mechanisms as the origins of centripetal and centrifugal forces (Fujita and Mori, 2005; Robert-Nicoud, 2005; Krugman, 2009). The overall effects of agglomeration externalities depend on the local range of these pecuniary externalities (i.e., proximity externalities). The effect is the same for intermediates and final goods as long as both suffer from transportation costs (Martin *et al.*, 2008; Combes *et al.*, 2008; Audretsch *et al.*, 2008).[93]

Pecuniary externalities operate via anonymous market interactions. If agents co-locate at a proximate distance, an anonymous market offers everything: providing agents with a large quantity and quality of inputs; efficient backward and forward market linkages;

involve a different process than the pure knowledge spillovers, the other form of R&D spillovers that Griliches noted."

92 See also Autant-Bernard and Massard (2007) and Capello (2007, 2009) for classifications.

93 The overall effect is interdependent and spurs agglomeration and the market size, which is denoted as a pecuniary externality or a pecuniary effect (Keilbach, 2000; Autant-Bernard and Massard, 2007; Capello, 2007).

retail firms in the neighborhood that reduce input costs and increase variety. Moreover, firms benefit from shared inputs in a local market, e.g., capital goods, intermediates and labor pooling (Johansson and Quigley, 2003; Duranton and Puga, 2004; Martin *et al.*, 2008). Thus, pecuniary effects can be regarded to enable firms to move to, or to move along, existing production frontiers (Neary, 2001; Harris, 2008; Capello, 2007, 2009). In comparison, non-pecuniary (technological) effects shift production possibility frontiers of firms and/or regions and countries (Romer, 1990b; Feldman, 1999, 2000; Harris, 2008). Furthermore, the concept of pecuniary externalities has been extended to the concepts of urbanization and localization economies, which represents a commonly used classification and focal point of empirical debates, especially in neo-Marshallian studies (see Feldman, 2000; Capello, 2007; DeGroot et al., 2009). These concepts first and foremost represent the "industry perspective" on agglomeration economies. However, the differentiation itself represents a highly discussed area of research which is additionally divided into "static" and "dynamic" externalities (Henderson, 2003a; Autant-Bernard and Massard, 2007; Duranton, 2008a). These externalities are also known as "efficiency" and "development" externalities (Johansson and Quigley, 2003; Johansson, 2005; Capello, 2007, 2009).[94] Table 2.2 shows the classification of pecuniary (efficiency) externalities into localization and urbanization economies and summarizes the different working channels. The concepts are separately reviewed in the following sections.

2.1.5.2. Localization Economies

"Localization economies" are assumed to usually take the form of Marshallian externalities. The labor productivity level in a certain industry is assumed to depend on the size of the industry and the specialization of the region (Dewhurst and McCann, 2007; Duranton, 2008a; Henderson, 2010). Rosenthal and Strange (2004) suggested that, for typical industries, doubling the local industry size leads to a 2-10% increase in the productivity level of employed workers. Moreover, doubling city size (i.e., the local scale) may also lead to a productivity increase in the same local industry, especially in high-technology industries. Dating back to Marshall's industrial districts argument, economies of scale due to high specialization, division of labor and increasing industry-wide output are associated with downward sloping average cost curves. Accordingly, industry concentration is assumed to promote external economies for agents, particularly at a proximate distance, and to have, first of all, static effects, i.e., input-costs, delivery costs (Johansson, 2005; DeGroot et al., 2009).[95]

94 According to Henderson (2003b, 29), *"[t]here appear to be two working interpretations of dynamic externalities. First is that either the history of economic activity in a location affects productivity levels or growth. So this could be past levels of own industry activity (employment) that generate a stock of local industry and location specific "trade secrets." The second set concerns the effect of "knowledge" (rather than information) spillovers on productivity levels. Knowledge is typically measured as non-industry specific, average education in the city. If average education in a city affects productivity it isn't clear this is a dynamic effect per se. It could be static in the sense that average education could simply enhance static productivity levels (but not on-going growth rates of productivity), but [...] that is sufficient to enhance overall urban scale and promote endogenous growth."*

95 See also Maggioni (2002), Autant-Bernard and Massard (2007), Burger et al. (2009) and Henderson (2010).

Table 2.2. Taxonomy of agglomeration economies

Localization economies (intra-industry)	**Urbanization economies (inter-industry)**
Accessibility to:	Accessibility to:
a specialized labor market	a diversified labor market
a high number of firms in the same industry	a high number of firms belonging to a diversified industry structure
specialized suppliers and service providers	a diversified market for industrial services and inputs
a highly specialized market for final goods (large market size)	a large market for diversified final goods
highly specialized R&D-departments and universities of intra-industry type	a diversified scientific environment with diversified universities and R&D-departments
scale economies from a single localized industry	scale economies from overall city size (also diversified production structure)
localization explains productivity level differentials of cities and regions	urbanization explains productivity level differentials of cities and regions

Source: own illustration.

A forerunning contribution to localization economies was made by Henderson (1974). He argued that an industry-specific externality in production decreases marginal costs of production depending on the level of industry output (i.e., efficiency externality). The distribution (and agglomeration) of firms is then determined by both positive industry externalities and negative effects from spatial concentration (e.g., commuting costs, increasing labor and capital costs). According to his results, cities show tendencies to specialize into industries. Thus, industries with large external economies tend to be predominantly concentrated in cities and urban areas.[96] Consequently, cities are considered to benefit significantly more from localization externalities than rural and peripheral areas (see also Henderson, 2010). Thus, a core-periphery distribution of industries and R&D-activity might represent a beneficial outcome.[97] The idea of such localization externalities (although the mentioned effects are pecuniary) has additionally been applied to innovative capacity in a spatial context (Feldman, 1999; van der Panne, 2004; Harris, 2008). However, such externalities are regarded as dynamic effects of co-location when they are related to growth in employment, innovations and new products (see section 2.1.6).

In a new economic geography context, Krugman and Venables (1996) and Aiginger and Pfaffermayr (2004) have argued that European economic integration might induce industry

96 Again, the contribution and differentiation of sources of externalities is rather controversial. Refer to Feldman (1999) and Glaeser (2000) for further details.

97 Ciccone and Hall (1996), among others, found that county employment densities are crucial in accounting for large differences in labor productivity across U.S. states.

agglomeration and increases in the degree of local specialization, bringing the European case closer to the spatial specialization pattern that economists and geographers identified in the United States. In this respect, they suggested that integration generally induces a geographic consolidation of industries at the national level and increasing regional specialization.[98]

To conclude, the local scale is a major part of the explanation why industrial activities agglomerate in cities and urban areas and why regions differ in terms of their productivity levels (Dewhurst and McCann, 2007). However, urban diseconomies (i.e., costs of co-location) dissipate agglomeration economies and explain why cities and localized industries show limitations in their overall size (Hoover, 1936; Henderson, 1974, 2003a; Duranton and Puga, 2004).[99]

2.1.5.3. Urbanization Economies

"Urbanization economies" are related to the size (urban scale) and industry structure of the city, region or agglomeration in order to explain varying levels in productivity (see table 2.2) (Hoover, 1936; Hoover and Giarrattani, 1999; Henderson, 2003). An obvious economic advantage of big cities is that they offer a large (and heterogenous) market affecting costs and prices (Johansson, 2005; Capello, 2007; Duranton, 2010).[100] Consequently, cities are considered to form and grow in order to exploit agglomeration economies that operate across all co-located activities/industries.

Hoover and Giarratani (1984, 73) discussed the advantages of urbanization as

> "[e]ssentially elements of a large urban agglomeration. Their presence, and the quality and variety of the services they offer, depend more on the size of the city than on the size of the local concentration of any of the activities they serve.[...] Economies generated by activities and services of this sort are external to any single-activity cluster, but they are internal to the urban area. There is a parallel to be drawn here to the relationship between a single-activity cluster and its constituent units. In that instance, economies were realized by the units as the size of the cluster increased; thus economies are internal to the cluster but external to the unit. In the case of urbanization economies, we recognize that economies accrue to constituent clusters as the size of the urban area increases. [...] Technological changes and enhancement of the mobility of labor and entrepreneurship explain why such local specialization has become increasingly rare. By contrast, external economies on the broader basis of urban size and diversity have remained a powerful locational force."

Similarly, Lucas (1993) discussed urbanization economies in an endogenous growth context, asserting that the main compelling reason for the growth of cities originates from

98 The specialization-localization debate will be reviewed in more detail in section 2.2.4 and is summarized in tables in the appendix.

99 See also World Bank (2009) and Henderson (2010).

100 Urbanization economies stem from effects external to the industry but internal to spatial units, such as cities or regions. For further discussions refer to Glaeser (2000), Feldman (2000), Maggioni (2002), Roos (2002b), Johansson (2005), Press (2006a), Autant-Bernard and Massard (2007), Capello (2007) and Dewhurst and McCann (2007).

the existence of increasing returns to scale that shift the productivity level.[101] Accordingly, urbanization economies can be empirically challenged by different working channels, depending on the research question, e.g., the employment structure, infrastructure, relative factor prizes, industry-structure, presence of industrial suppliers and service providers (Johansson, 2005; de Groot *et al.*, 2009; Puga, 2010).

However, it has also been argued that cities, metropolises and urban regions may reach absolute population levels and scales that increase external diseconomies of congestion that fully (or even more than) compensate positive agglomeration economies (Duranton and Puga, 2004; Martin *et al.*, 2008; World Bank, 2009).[102]

Finally, Partridge and Rickman (1999) and Burger *et al.* (2008) argued that static externalities from urbanization are different from dynamic urban externalities, the so-called Jacobs externalities (see section 2.1.6.3), as the latter explicitly account for employment growth, productivity growth and the effects from knowledge spillovers on innovative capacities (product development and new routines), whereas the static urbanization externality concept in general focuses on the local scale of production and existing productivity differentials, and the positive effects from population density, input supply structures and local market size on prices and costs. In light of the still prevalent debate and area of regular conflict, recent approaches link the Jacobs externality concept to the process of inter-industry knowledge diffusion and the effects of innovation externalities on innovative capacity (Audretsch and Feldman, 2004). The approach is thus different from the 1970s static view on agglomeration economies and the explanation of persistent productivity differentials (Partridge and Rickman, 1999; Johansson, 2005; Capello, 2009).[103] Similarly, Capello recently argued that especially neo-Marshallian studies identified regions (and space in general) as the origin of "dynamic" external economies which emerge as positive effects from co-location and affect the firm's productive and innovation efficiency (i.e., innovative capacity) (Capello, 2007, 185; Capello, 2009). Accordingly, the effects differ from static gains in agglomerations.

2.1.5.4. A Taxonomy of Urbanization and Localization Economies

Table 2.3 summarizes pecuniary (intra-market) externalities that originate from spatial proximity in an urbanization and localization economies context, where the emphasis is on

---

[101] See Roos (2002b), Press (2006a), Capello (2007, 2009) for an overview.

[102] Higher GDP per capita growth rates and other forms of agglomeration economies only dominate up to metropolitan size of maximum 6 to 7 million citizens, which resembles the well-known inverted U-shaped relationship (OECD, 2006, 2009a,f; World Bank, 2009) (see section 5.4). Another crucial debate is about the privatization of benefits from agglomeration and the socialization of associated costs that stem from concentration (OECD, 2009b,a,f). See also Capello (2007), World Bank (2009) or Henderson (2010).

[103] Partridge and Rickman (1999, 319) argued that *"[s]everal empirical regional studies related to geographic concentration of economic activity and economic spillovers emphasize their relationship to employment growth, only indirectly testing the externality-productivity relationship (e.g., Glaeser et al. 1992; Henderson, Kuncoro, and Turner 1995; Partridge and Rickman 1996; Henderson 1997). Also, studies of regional productivity differences typically focus on static urbanization and localization economies and not on dynamic externality effects emphasized in the endogenous growth literature (e.g., Moomaw 1983, 1986)."*

prices, costs, profits and productivity levels. More generally, the demand-externality is related to the size of the market, which is central in the new economic geography (Krugman, 1991). To conclude, localization and urbanization economies emerged as a conceptualization in the mid 1970s which represents the conventional industry dimension of agglomeration economies (Feldman, 2000; Capello, 2007, 2009). The focus of analysis is solely on the question whether agglomeration economies are related to the scale of a single (and specialized) industry or to the cross-fertilization enhanced by diversity and the scale of other industries (Henderson, 2003a). Consequently, the concept centers the exploitation of indivisibilities within either a specialized or diversified industry environment (Feldman, 1999, 2000; Rosenthal and Strange, 2001; Johansson, 2005). Thus, the presence and strength of the mentioned externalities is related to the emergence and stability of core-periphery structures.

Table 2.3. Pecuniary externalities

Externality type	Transaction type	Mechanism
Demand	Intra-regional demand externality. "Home market effect."	A large local demand makes it possible for firms to exploit scale economies and hence supply commodities to households at a lower price and with a greater variety.
Vertical	Downstream/delivery-cost externality. Localization and urbanization economies.	Firms can offer inputs/products with lower transaction costs and (potentially) at a lower price due to physical proximity (proximity externality).
	Upstream input-cost externality. Localization and urbanization economies.	Supplier firms in an industry provide inputs with lower transaction costs and (potentially) at a lower price due to physical proximity (proximity externality).

Source: illustration based on Johansson (2005, 119), Johansson and Quigley (2003) and Capello (2007, 2009).

2.1.5.5. Core-Periphery Structures and Endogenous Location

2.1.5.5.1. The Origins of the New Economic Geography

The new economic geography corresponds to the proposed classification of agglomeration economies (see section 2.1.4.3) and represents a formalization of pecuniary externalities that induce core-periphery structures. Paul Krugman is one of the main contributors to geographical economics in the 1990s who has been working for re-establishing spatial aspects as a pivotal factor in economic theorizing and empirical research (Krugman, 1991, 1992, 2009; Fujita and Krugman, 2003; Fujita and Mori, 2005).[104]

According to Krugman (1995), a salient feature of early trade and geography models is the assumption that countries specialize their production on the locally abundant factor, e.g., natural resource advantages (e.g., the Heckscher-Ohlin framework within neoclassical trade theory).[105] It has been argued by Krugman and colleagues that endogenous spatial issues, centripetal forces and circular causalities are rather absent in early trade and geography models (Head and Mayer, 2004; Krugman, 2009; Neary, 2009). As Krugman (2009, 567) provocatively noted:

> "[W]hy was geography ignored by trade theorists? A large part of the explanation is the obvious centrality of increasing returns to geographical patterns: nobody really thinks that Silicon Valley owes its existence to exogenously given factors of production or Ricardian comparative advantage [although] God made the Santa Clara Valley for apricots, not semiconductors."

A central feature of the NEG framework is that it completely abstracts from physical geography, i.e., "first-nature causes" (see chapter 2, section 2.1.2) (Ottaviano and Thisse, 2000; Capello, 2007; Thisse, 2011).[106] The crucial difference between the new economic geography approach and the neoclassical models is the overall relevance of scale in production (Roos, 2002b; Crafts and Venables, 2003; Capello, 2007). Neoclassical (trade) models are solely concerned with relative terms, suppressing scale and size; e.g., consumers' choices for saving and consumption, firms' decisions of production structures, and wage-setting are all determined at the margin. According to Krugman (1992), among others, the outcome of such processes are unaltered in an economy with 10, 10,000 or 10,000,000 individuals. Regarding this aspects, Krugman (1992, 14) argued that

> "[t]he basic story of geographic concentration [...] relies on the interaction of increasing returns, transportation costs, and demand. Given sufficiently strong economies of scale, each manufacturer wants to serve the national market from a single location with large local demand. But local demand will be large precisely where the majority

[104] Krugman was awarded with the Nobel Price in Economics in 2008 for his contributions to trade theory and economic geography. For an appraisal see Nobel Prize Committee (2008a), Nobel Prize Committee (2008b), Fujita and Thisse (2009), Feenstra (2009), Neary (2009), Brakman and Garretsen (2009).

[105] Brakman and Garretsen (2009, 2) defined the economic novelty of the NEG literature as follows: *"[T]he subsequent NEG literature can in fact be seen as belonging to a much more extensive (and older) literature in regional economics or even economic geography at large, where spatial interdependencies are at the heart of the analysis. The performance of a region depends crucially on the developments in and characteristics of neighboring regions. Regions are therefore not freely floating islands in NEG. This non-trivial role of spatial linkages amounts to saying that it is above all between location economic geography that matters in (old and) NEG."* See also Ottaviano and Thisse (2000, 3).

[106] See also Krugman (1991), Krugman (1995) and Krugman (2000). Fujita *et al.* (2001, 4) mentioned the importance of incorporating the spatial dimension into economic theory and empirical research. They argued that *"[t]he defining issue of economic geography is the need to explain concentrations of population and of economic activity: the distinction between manufacturing belt and farm belt, the existence of cities, the role of industry clusters. Broadly speaking, all these concentrations form and survive because of some form of agglomeration economies, in which spatial concentration itself creates the favorable economic environment that supports further or continued concentration."*

> of manufacturers choose to locate. Thus there is a circularity that tends to keep a manufacturing belt in existence once it is established."

Thus, scale effects (i.e., increasing returns) in production are considered to significantly matter in the new economic geography and a core-periphery distribution might be the outcome of relocation activities and large distributional effects via self-feeding and self-enforcing mechanisms.[107] Krugman's model is based on the Dixit and Stiglitz (1977) contribution of product differentiation but it also offers a formalization of Myrdal's circular and cumulative causation (Krugman, 1995; Crafts and Venables, 2003; Capello, 2007).[108] Nevertheless, NEG is considered to go beyond forerunning contributions because agglomeration is not the only possible outcome.[109] Krugman (1991) also distinguished the new economic geography from previous land-rent models that were based on, inter alia, von Thünen (1966) and Alonso (1964), among others (Fujita and Krugman, 2003; Fujita, 2010; Thisse, 2011).[110] NEG adherents generally argued that the work of classical contributors, e.g., Johann Heinrich von Thünen (Thünen, 1966), are unsuited to explain the emergence and breakdown of core-periphery structures (and intra-industry trade) due to the lack of agglomeration economies (Krugman, 1995; Roos, 2002b; Combes *et al.*, 2008). In the majority of land-use models, the existence of a core region (or central market) is exogenous and the main attention is on accessibility and transportation costs, which makes this optimization criterion central to urban economics (Capello, 2007; Fujita, 2010; Thisse, 2011). The NEG framework and model alternatives are related to several agglomeration economies mentioned in Marshall's *Principles of Economics* (see section 2.1.4.2). A detailed critical discussion, theoretical subordination and comparison of early location theories and contributions is, however, beyond the scope of this project.[111]

As the NEG frameworks are mostly identical, except the origins of centripetal and centrifugal forces, the subsequent section briefly reviews the new economic geography framework in its simplest form (i.e., the Krugman (1991) model). Afterwards, a brief summary of al-

107 Krugman (1995, 36) has argued that *"[i]n order to talk even halfway about economic geography it is necessary to invoke the role of increasing returns in some form."*. Similarly, Weber (1922, 123) already discussed agglomeration effects and locational advantage similar to Krugman's NEG as a *"[V]orteil, also eine Verbilligung der Produktion oder des Absatzes, die sich daraus ergibt, dass die Produktion in einer bestimmten Masse an einem Platz vereinigt vorgenommen wird."*

108 Myrdal (1957) addressed circular causation similar to Hirschman (1958). Circular causation and cumulative effects are considered to lead to increasing superiority of countries that already have superior productivity and a high level of income, while inferior countries will suffer from inferior income levels. Myrdal described some backwash effects that promote inequality and emerge from internal and external economies (economies of scale and growth of knowledge by innovation). See Combes *et al.* (2008).

109 Essential for Krugman's framework is that a fraction of the mobile income is spent in the region in which it is generated. NEG approaches are said to be built upon qualitative ideas of Perroux's (1955) "growth poles," on Myrdal's (1957) contributions to "circular and cumulative causation," and finally on Hirschman's (1958) concept of "forward and backward linkages" (Krugman, 1995; Crafts and Venables, 2003; Capello, 2007; OECD, 2009a).

110 See also Krugman (1995), Roos (2002b), Robert-Nicoud (2005), Fujita and Mori (2005), Capello (2007), Combes *et al.* (2008), OECD (2009a) and Dauth (2010).

111 The interested reader should consider, e.g., Roos (2002b), Roos (2004), Capello (2007), Cruz and Teixeira (2007), Combes *et al.* (2008), OECD (2009a) and Thisse (2011) for a comprehensive overview and discussion of agglomeration theories.

ternative core-periphery models will follow with special emphasis regarding modifications, extensions and conclusions.[112]

2.1.5.5.2. Industry Agglomeration, Core-Periphery and Footloose Labor

Krugman's (1991) new economic geography (NEG) framework is regarded as the starting point of a new generation of spatial economic models (Krugman, 1991, 1995, 2009; Capello, 2007; Combes *et al.*, 2008).[113] Increasing returns, transportation costs and migratory movements are the pivotal factors which determine the dynamics of centripetal and centrifugal forces and the emergence of core-periphery structures.[114] The core-periphery model instruments the trade-off between positive pecuniary externalities (i.e., agglomeration economies, see section 2.1.5) and transportation costs.[115]

As skilled and unskilled labor are the only production factors, which are assumed to be partly mobile, the modeling equilibrium essentially depends on the migration behavior of workers. The model simplifies by excluding capital and intermediates. It consists of two sectors; an agricultural and manufacturing sector. Workers in the agricultural sector do not feature inter-regional or at least inter-sectoral mobility (i.e., the centrifugal force). In opposition, workers within the manufacturing sector (and their expenditures) are inter-regionally mobile, which represents the centripetal force. The representative agent is modeled to follow a CES utility function, which exhibits a two-stage process of expenditure allocation, i.e., expenditures for the agricultural (homogeneous) good and expenditures for varieties of the manufacturing sector. The production side is defined by constant returns to scale in the agricultural sector and increasing returns to scale in the manufacturing sector (Krugman, 1991, 1995; Capello, 2007; Combes *et al.*, 2008).

Regarding endowments, every region has a fixed amount of labor in the agricultural and manufacturing sector. The traditional good underlies no transportation costs (i.e., the channel for factor prize equalization), whereas manufacturing goods exhibit costs of transportation. The equilibrium location of workers/ consumers and firms is determined by several forces that strongly affect the dynamics of the model in terms of stability and adjustment mechanisms (Krugman, 1991, 1995; Capello, 2007; Combes *et al.*, 2008).

Firms optimize economies of scale in production by their location decision, which represents the incentive to relocate production to the larger regional market; i.e., the "market-size-effect" or "home market effect." The market expansion affects the local profits and represents a strong incentive for firms to relocate production according to absolute terms

---

112 The following section only summarizes the main conclusions. For comprehensive overviews and formal representations of NEG models refer to, e.g., Roos (2002b), Fujita and Krugman (2003), Robert-Nicoud (2005), Fujita and Mori (2005), Capello (2007), Eckey (2008), Combes *et al.* (2008) and Thisse (2011).

113 For an overview refer to Roos (2002b), Litzenberger (2007), Eckey (2008) and OECD (2009a).

114 In general, the NEG is widely classified as an extension of the new trade theory, although Krugman personally defined the NEG as a general framework which includes the new trade theory as a special case.

115 Krugman (1995, 62) finally stated that *"[t]he clustering of production that results from this dynamic process can be seen as the consequence of a kind of pecuniary external economy, not really inconsistent with Marshall's description."*

(Fujita and Krugman, 2003; Fujita and Mori, 2005; Krugman, 2009). A firm's decision to relocate production induces two processes that essentially affect regional disparities. First, there is a competition effect: the market entry of firms squeezes the market share and profits of existing firms in the local market. Second, the spatial distribution is affected by a demand or market-size effect: the increasing size of the local market affects the profits of firms and their local labor demand. The increasing number of varieties and labor demand bring about higher local wages and attract mobile skilled workers; thus, the circular causality induces an additional increase in local expenditures via migratory movements. As skilled workers are the only mobile factor of production, inter-regional movements induce expenditure shifting, followed by production shifting. The higher the number of consumers in a location, the more firms have to employ an increasing workforce to supply goods in the local market. The more varieties are produced, the higher is the level of real income and the more workers are attracted to the region. Accordingly, demand shifting induces production shifting and the relocation of manufacturing goods induces again demand and expenditure shifting (centripetal force) but also higher competition in the labor and goods market in the growing core region (centrifugal force). A perpetuum mobile is clearly visible (Krugman, 1991, 1992, 2009; Capello, 2007; Combes *et al.*, 2008; OECD, 2009a). However, the fixed amount of immobile agricultural workers functions as a strong centrifugal force against the agglomeration process because immobile workers represent an immobile source of demand for manufactured goods in the periphery. Low levels of transportation costs will, however, not stop the relocation of skilled workers and the emergence of an agglomerated manufacturing industry (Krugman, 1991, 1992, 2009; Henderson, 2003a; OECD, 2009a).

Nevertheless, the emergence of industrial core-periphery structures depends on the net effect, meaning that the centripetal force (demand effect) has to surpass the centrifugal force (competition effect). The strength of the competition effect depends positively on the substitution elasticity among the goods produced by manufacturing firms and the costs of transportation. The intensity of the demand effect depends on the realized economies of scale which increase profits and the income share spent on manufacturing goods (Krugman, 1991, 1992, 1995, 2009; Capello, 2007). Accordingly, the original core-periphery model is central in explaining the formation of core-periphery structures detached from endowments.

To conclude, the stability of a core-periphery pattern is highly influenced by transportation costs and the presence of pecuniary externalities. The demand effect always exceeds the competition effect in case that (i) varieties are difficult to substitute, (ii) scale economies are intense, (iii) the share of income spent by consumers on manufacturing goods is large and (iv) transport costs are at a modest level or decreasing due to integration (Krugman, 1991, 1992, 1995, 2009). Accordingly, the strength of the competition effect and demand effect are considered to change the geographic distribution of industries and the size of local markets in a meaningful way (Krugman, 1991, 1992, 2009; Capello, 2007). Small variations of the initial distribution of labor crucially determine the spatial distribution of the manufacturing industry and the expenditures of consumers (function of initial conditions or historical accident). Without the above mentioned assumption of migratory movements of manufacturing workers, the core-periphery model would be unable to produce core-periphery structures (Krugman, 1991; Roos, 2002b; Capello, 2007). Thus, the implementation of inter-regional mobility of workers is a key factor in explaining spatial expenditure shifting and thus circular and cumulative causation known from early cumulative development models (Krugman, 1992, 2009; OECD, 2009a; Thisse, 2011).[116] In retrospect, new economic geography is considered to have benefited considerably from several past workhorse contributions in economic theory, economic geography and regional science (Fujita and Krugman, 2003; Crafts and Venables, 2003; Combes *et al.*, 2008).[117]

For comparison purpose, the main findings of selected alternative NEG models, which offer alternative centripetal forces, are briefly reviewed in the subsequent section.[118]

2.1.5.5.3. Alternative Core-Periphery Models

Krugman and Venables (1995b) introduced a model that puts aside migratory movements of workers; it consists of two regions that are identical in endowments, preferences and technology. The manufacturing sector produces final goods and intermediates under increasing returns to scale technology. In case of high transportation costs, the manufacturing industry is equally distributed across the regions. In case of low transportation costs (below a certain threshold level), the region with the initially larger manufacturing share will induce a relocation process and attract producers due to strong forward and backward industrial linkages (demand and cost linkages). First, final good producers benefit from stronger industry concentration of intermediate producers, which induces forward linkages (i.e., cost linkages). Second, producers of intermediates will prefer to produce at a proximate distance to large final good producers, which induces backward-/demand-linkages (i.e., the centripetal force). Real income in the core region is increasing due to these forward and backward linkages. However, in case of falling transportation costs, the wage differential between the periphery and the core region will induce relocation of firms back into the periphery. The model is pivotal in explaining input-output structures in the manufacturing industry.[119]

Venables (1996) introduced an alternative framework to the Krugman-type core-periphery model without migratory movements. The model incorporates pecuniary externalities which originate from industry input-output linkages that induce cost effects. The model consist of three sectors (in opposition to two sectors in Krugman and Venables (1995b)). One sector produces a tradable good (perfectly competitive sector); the other two sectors

---

116 Furthermore, Krugman (1991, 497) argued that *"[i]n an economy characterized by high transportation costs, a small share of footloose manufacturing or low economies of scale, the distribution of manufacturing production will be determined by the distribution of the primary stratum of peasants. With lower transportation costs, a higher manufacturing share or stronger economies of scale, circular causation sets in, and manufacturing will concentrate in whichever region gets a head start."* Refer also to Nobel Prize Committee (2008a,b).

117 See also Capello (2007), OECD (2009a) and Neary (2009). As Fujita and Krugman (2003, 142) have concluded: *"Dixit-Stiglitz, icebergs, evolution and the computer. Yet the slogan captures the essence of the intellectual tricks that we and other new economic geography theorists have used in order to cope with the technical difficulties involved in trying to deal with the subject. Everyone recognizes that these are strategic simplifications, which is to say, intellectual cheap tricks; but they do allow us to get past the technical issues and tell the stories about the real economics."*

118 For an overview refer to OECD (2009a).

119 For a comprehensive overview see Keilbach (2000), Roos (2002b), Litzenberger (2007), Combes *et al.* (2008) and OECD (2009a).

are vertically linked and show a monopolistically competitive structure and one sector produces an intermediate good. The industries are located in both regions and the firms supply their output to both markets. The level of linkages and transportation costs determines the production decision of firms. In the case of high transportation costs, firms have an incentive to locate close to consumers and thus to produce in both locations. In the case of low transportation costs, firms also prefer to produce in both locations, which induces decreasing regional disparities since factor prices are equal in both locations. In the case of intermediate transportation costs, the model shows emerging clustering forces which give rise to multiple equilibria, determined by factor price differences. Some industries show symmetric distribution in response to regional factor price differences; other industries, however, agglomerate. To summarize, the cost effect in the model is associated with an increase in firms' profits through an expansion of the local market for intermediates. This centripetal force works against the local competition effect (centrifugal force). If firms use the same intermediates, the model of Venables predicts the following: (i) a decreasing intermediate good price for downstream firms; (ii) an increasing market size for upstream ones. The contribution is central in explaining the emergence of an intermediate industry and the strength of pecuniary externalities.[120]

Krugman and Venables (1996) modified Venables (1996). The authors placed emphasis on the process of European integration. Their model includes two industries in two regions. Final and intermediate goods are produced in both industries and intermediates are used in the production process. The production technology is characterized by monopolistic competition. Migratory movements of workers are absent and transportation costs exist for manufacturing goods. The model emphasizes the dynamics of economic integration between several regions, each containing several industries. In case of high transportation costs, the regions will generally maintain the full range of industries. Backward and forward linkages are not strong enough to lead to core-periphery structures. In case of intermediate transportation costs, industry agglomeration can only be observed if the initial distribution of industries is skewed. In case of low transportation costs, regions with a strong initial (exogenous) industry endowment benefit from a locational advantage that evolves due to forward and backward linkages (cumulative circular causation). In this case, each region concentrates onto a single industry.[121]

2.1.5.5.4. Critical Remarks and Discussion

The NEG literature contains a large number of alternative frameworks which explain the spatial distribution of activities at different geographical levels, e.g., international specialization, the distribution of employment and productivity at the national, regional and city level (Neary, 2001; Roos, 2002b; Fujita and Krugman, 2003).[122] The impressive amount of models is based upon heterogeneous (alternative) ways in which inter-firm relations generate externalities that induce centripetal and centrifugal forces. In some cases the externalities originate from the demand-side, in others from input-output linkages at the supply-

[120] For an overview refer to Roos (2002b), Litzenberger (2007) and OECD (2009a).
[121] Refer also to Keilbach (2000), Roos (2002b), Litzenberger (2007), Combes *et al.* (2008) and OECD (2009a).
[122] See alsoCapello (2007), Combes *et al.* (2008) and Neary (2009).

side. Nevertheless, all these models implement pecuniary externalities.[123] Accordingly, the models solely differ in the way economists and geographers allow for varying modes of mobility of people, capital goods or intermediates. The purpose of first-generation NEG models was to explain regional specialization, agglomeration (i.e., industrialization) and the distribution of industries (and agents) across regions, depending on spatial distance, market-size (pecuniary) effects and costs of transportation.[124] Non-pecuniary externalities, such as knowledge flows or externalities (section 2.1.7), have not played any role.[125]

Furthermore, a salient feature of all new economic geography models is that the location choice of firms is solely determined by profit maximization behavior whereas the location decision of households crucially depends on utility maximization. The activities of agents are interrelated and induce cumulative circular developments that give rise to centripetal and centrifugal forces in line with antecedent contributions; i.e., Myrdal's virtuous circle of cumulative development and Kaldor's formalized model of cumulative circular causation (Krugman, 1995, 2009, 2010; Capello, 2007, 228, 236; Thisse, 2011). The basic argument, according to early new economic geography adherents, was to develop a general theoretical framework, that also includes aspects of trade theory, with special focus on pecuniary (and proximity) externalities, which enter the system via market mechanisms depending on prices (Krugman, 1991, 1995; Capello, 2007). Opposed to technological externalities, market-led mechanisms do not affect utility or production directly (Ottaviano and Thisse, 2001; Roos, 2002b; Harris, 2008).[126]

Although there exists a meaningful amount of modifications of the new economic geography, the majority of (older) contributions have not considered knowledge spillovers as a major driver of agglomeration (Krugman, 1995, 2009; Martin, 1999; Capello, 2007). Krugman was highly pessimistic with respect to centripetal forces stemming from knowledge spillovers.[127] Krugman (1992, 54) argued that

> "[k]nowledge flows are invisible; they leave no paper trail by which they may be measured and tracked, and there is nothing to prevent the theorist from assuming anything about them that she likes. So while I am sure that true technological spillovers play an important role in the localization of some industries, one should not assume that this is the typical reason - even in the high technology industries themselves."

[123] The models which include R&D are briefly discussed in section 2.1.6.

[124] In contrast to first-nature causes that refer to the concept of comparative advantage (i.e., natural advantage, resources, endowments, infrastructure, climate, past location choice), the NEG models especially use second-nature causes of agglomeration that induce cumulative causations due to pecuniary externalities.

[125] For a critical debate on spillovers refer to Fujita and Thisse (1997), Krugman (2000), Fujita and Krugman (2003), Robert-Nicoud (2005) and Krugman (2009, 2011).

[126] Krugman (1995, 52) argued that *"[w]hile it has been possible to make the sources of agglomeration safe for neoclassical economics by assuming that they are pure technological externalities, this strategic evasion has been costly in terms of both credibility and researchability. [Consequently,] you have no deeper structure to examine, no way to relate agglomeration to more micro-level features of the economy."*

[127] In this context, the spatial share of innovative firms and the effects of firm size on invention and productivity are essential in terms of intensity and spatial scope of MAR, Porter or Jacobs externalities and the cluster life cycle. Unfortunately, firm size is still not incorporated in the early NEG models.

Therefore, the new economic geography was mainly focusing on the trade-off between pecuniary externalities (centripetal force) and transport costs (centrifugal force). In this respect, first-generation NEG models consciously suppressed technological externalities. However, the influence and significance of knowledge externalities in the context of knowledge-intensive industries became increasingly issued (Fujita and Thisse, 1996; Audretsch and Feldman, 1996); especially by economic geographers, geography of innovation adherents and researchers in an innovation system tradition (Duranton and Rodríguez-Pose, 2005; Behrens and Thisse, 2007; Rodríguez-Pose, 2010). Additionally, Krugman (2011) has put his 1990 statement into perspective.[128]

Another shortcoming of the presented new economic geography framework(s) is that social and relational proximity, in the context of knowledge-intensive industries, do not perform a central role in the conceptualization of agglomeration economies. This matter of fact distinguishes the NEG from the locational advantages in innovative milieus, network approaches and concepts in an evolutionary economics tradition (see section 2.1.7.3). Accordingly, NEG solely applies pecuniary externalities (see section 2.1.5) but not technological externalities (see section 2.1.6). A final critique concerns the lack of an R&D industry in early NEG models. As R&D activity is not explicitly modeled, regional disparities of research activities have to be considered to represent a reflection of the spatial distribution of the manufacturing industry (i.e., the distribution of the intermediate and final good industry). Thus, it has to be assumed that the distribution of R&D solely follows the spatial distribution (relocation) of the manufacturing industry. Accordingly, the framework clearly suppresses R&D activities and knowledge-intensive tasks, which seems to reduce its applicability to knowledge-intensive industries (Krugman, 2009, 2011). To address this issue in a European context, the distribution of knowledge-intensive tasks is measured directly by analyzing the spatial concentration of European research (and patenting) activity at the regional level (see chapter 3).

2.1.6. Industry and Research Clustering and Innovation Externalities

2.1.6.1. Non-Pecuniary Externalities

The agglomeration economies, which have been described in the previous sections, examine the role of space as a promotor of locational advantage that enters the system by means of a more efficient use of resources and lower transaction and production costs, which induce higher levels of productivity and profits at the firm-level. However, spatial proximity not only improves static efficiency of production. Economists and geographers have additionally stressed the importance of non-pecuniary effects, i.e., technological externalities, in the context of research clustering, R&D networks and co-patenting activity (Feldman and Audretsch, 1999; Audretsch and Feldman, 2004; Johansson and Quigley, 2003). This consideration predominantly concerns the relationship between spatial proximity and innovative and creative capacities of the firm and individuals, i.e., innovation externalities (Capello, 2007, 193; Capello, 2009).

[128] Similarly, Fujita and Thisse (1996, 345) argued that *"[a]n economic agglomeration is created through both technological and pecuniary externalities, often working together."*

A seminal contributor to the technological externalities debate was Scitovsky (1954). He concluded in the 1950s that non-pecuniary externalities apply when agents are interdependent.[129] However, interdependence between agents does not necessarily have to occur via prices, i.e., via anonymous market transactions.[130] His main focal point was on the question of non-rivalry, non-excludability and compensation. A few decades later, Griliches (1992b, 36) has related non-pecuniary externalities to R&D activities and knowledge spillovers as

> "[w]orking on similar things and hence benefiting much from each others research. [...] True knowledge spillover are ideas borrowed by the research teams of industry *i* from the research results of industry *j*. [...] To measure them directly in some fashion, one has to assume either that their benefits are localized in a particular industry or range of products or that there are other ways of identifying the relevant channels of influence, that one can detect the path of the spillovers in the sands of data."

Accordingly, technological externalities are considered to influence the economic efficiency in terms of new routines or innovative capacity and innovation output of the firm and to increase product diversity (Johansson, 2005; Capello, 2007).

Moreover, it is frequently argued that non-pecuniary externalities represent effects disembodied from capital goods, new products, intermediates and service inputs and to be decoupled from direct input-output linkages (Fujita and Thisse, 1996; Duranton and Puga, 2004; Capello, 2009). However, the technological externality approach (i.e., knowledge spillover approach) has been heavily criticized as it implements a "black box" of unexplained technological progress into the production system (Fujita and Krugman, 2003; Breschi and Lissoni, 2001a; Breschi *et al.*, 2005).[131]

With respect to regional innovation systems (Cooke *et al.*, 1997; Doloreux and Parto, 2005) and research clusters, an essential factor is R&D activity, as it has a twofold effect within the region. The first effect has a direct nature and leads to new blueprints/patents and new products (see also sections 2.1.6.6 and 2.1.6.7). The benefits are then appropriated by the inventor who has, under the assumption of existence of an intellectual property protection system, a temporary monopoly position and rents. The second effect comes from knowledge codification (i.e., the technical documentation/blueprint) in the patent application process. This effect has an indirect nature in the sense that this kind of knowledge is available for competitors under specific circumstances. They can accumulate the knowledge which positively influences their innovative capacity (Capello, 2007, 189). However, due to the property right protection competitors cannot commercialize the same pieces of knowledge

---

129 According to Scitovsky (1954, 144), non-pecuniary externalities originate from *"[...] inventions that facilitate production and become available to producers without charge."*

130 Scitovsky, among others, has differentiated between rent spillovers (via traded goods) and technological spillovers (no compensation). Rent spillovers originate from a market transaction, whereas technological externalities (pure knowledge spillovers) are assumed to occur outside the market, e.g., knowledge spillovers from just being there. For older discussions see Griliches (1979), Grossman and Helpman (1991b), Coe and Helpman (1995), Coe *et al.* (1997), Verspagen (1997) and Keilbach (2000).

131 Ottaviano and Thisse (2001, 160), for example, argued that *"[t]echnological externalities are 'black boxes', that is, 'reduced forms' that capture the crucial role of complex non market institutions whose role and importance are strongly stressed by geographers and urban planners."*

as the patent system generally provides a legal mechanism to enforce excludability.[132] Nevertheless, the overall effect is a technological externality and an upgrading of the region-wide, perhaps industry-wide, stock of knowledge and a positive effect on the innovative capacity (i.e., innovation externality) (Johansson, 2005; Capello, 2007).[133]

Directly related to the previous points is the debate whether knowledge always fulfills the attributes of non-rivalry and non-excludability (see, e.g., Romer, 1990b). The public good character has been heavily criticized (Lissoni, 2001; Breschi and Lissoni, 2001b,a; Krugman, 1995).[134] Several research streams have implemented the idea of externalities but modified to localized knowledge spillovers, localized networks and tacit knowledge (see section 2.1.7). In fact, the non-rivalry assumption of knowledge mainly works as perpetual-motion in endogenous growth and economic geography models (see sections 2.1.6.6 and 2.1.6.7). Endogenous growth models which are heavily built upon technological externalities normally lead to hypotheses towards economy wide under-investments in R&D and knowledge production (see section 2.1.6.6).[135] Due to the effects of R&D externalities on aggregate growth, productivity and innovation output, the degree of rivalry and excludability crucially determines normative conclusions and policy recommendations. These debates are essentially dependent on the spatial level of analysis as externalities are determined by strong distance decay effects, networks and industry structures (see also sections 2.1.7 and 2.2).

2.1.6.2. Marshall-Arrow-Romer Externalities and Specialized Clusters

A by-product of the discussions on agglomeration economies, associated with the knowledge spillover approach, is the introduction of the "Marshall-Arrow-Romer (MAR) externalities" concept (Audretsch, 1998; Partridge and Rickman, 1999; Carlino *et al.*, 2001).[136] It is argued that local firms benefit from externalities that originate from a high concentration of firms of the same industry (localization). These externalities can be regarded as a dynamic type of localization economies (dynamic agglomeration economies) which are linked to inventive capacity and innovation output of firms in the same industry (Feldman, 2000;

[132] Non-pecuniary effects (externalities) determine the individual utility or production function directly (Combes et al., 2008). This aspect is a crucial determinant in the new economic geography variants of Martin and Ottaviano (1999), Baldwin and Forslid (2000a), Baldwin *et al.* (2001b), Baldwin and Martin (2004) as will be discussed in detail in section 2.1.6.7.

[133] According to the mainstream literature, the diffusion of knowledge in such models is not accompanied by (full) monetary compensation.

[134] Following Romer, non-rivalry means that different firms (or regional units) can take advantage from new pieces of knowledge without diminishing the flow of knowledge for competitors (Romer, 1986, 1990b). Even if firm A diminishes the long-term economic benefit from the piece of knowledge for competitors B, C, and D, all of them gain by absorbing (technological) knowledge/blueprints without compensation.

[135] Griliches (1992b, 43) has put special emphasis on R&D spillovers and has concluded that *"[...] there has been a significant number of reasonably well done studies all pointing in the same direction: R&D spillovers are present, their magnitude may be quite large, and social rates of return remain significantly above private rates."*

[136] The MAR concept is considered to originate from contributions of Marshall [1890](1920), Arrow (1962) and Romer (1986). Refer also to the remarks of Feldman (1999), Keilbach (2000), Combes (2000b), Paci and Usai (2000a), van der Panne (2004), van der Panne and van Beers (2006), Capello (2007) and de Groot *et al.* (2009).

Johansson, 2005; Capello, 2007).[137] Accordingly, MAR-externalities imply that the spatial concentration of firms of a specific sector promotes growth rates of innovation output (and implicitly industry employment and productivity growth) (van der Panne, 2004; Johansson, 2005).[138]

In Henderson (1974), localization economies in cities are considered as a pure form of the above presented Marshallian externalities; industry-specific productivity of workers is assumed to increase with the industry employment share. Moreover, a relatively small technological distance between individuals and firms implies low barriers for knowledge spillovers and is seen as a condition for sustained growth (innovation externality) (Greunz, 2003a; van der Panne, 2004; Döring and Schnellenbach, 2006). This would lead to the testable hypothesis, that regions which are characterized by higher sector-/industry-specific localization of firms, and thus similar production technologies, should ceteris paribus tend to have higher income and productivity growth rates and/or innovation output than regions with lower intra-industry localization (Griliches, 1979; Boschma and Frenken, 2009a).[139] Accordingly, Marshall-Arrow-Romer externalities are sometimes linked to knowledge spillovers and innovative capacity and sometimes to productivity or employment growth (Beaudry and Schiffauerova, 2009).[140]

Similarly, Feldman (2000) concluded that firms, if they belong to the same (or at least similar) industry, in which complementary assets are essential, could realize greater gains (i.e., growth) in productivity.[141] Moreover, she argued that MAR-externalities are generally associated with strong distance decay effects (proximity externalities).

A final consideration in this context concerns the industry and cluster life cycle and the existence of high- and low-tech regions (Audretsch and Feldman, 1996; Audretsch, 1998; Malecki, 2010). Audretsch and Feldman (1996, 254) have matched the concept of knowledge transfer with the stage of the (spatially clustered) industry under observation. They concluded that

> "[p]erhaps most striking is the finding that during the mature and declining stages of the life cycle increases in the geographic concentration of production tend to lead to greater and not less dispersion of innovative activity. It may be that new ideas need new space, at least during the mature and declining stages of the industry life cycle. In any case, the positive agglomeration effects during the early stages of the industry life cycle apparently are less important during the latter life cycle stages."

137 Empirical studies in the neo-Marshallian tradition also use productivity growth or employment growth as proxies for such dynamic agglomeration economies (see also Roos, 2002; Harris, 2008; Neffke et al., 2009).

138 To repeat a point made earlier, static localization economies solely explain disparities in (regional) productivity levels and the effects on prices and costs via market transactions (Glaeser et al., 1992; Audretsch and Feldman, 1994, 1999; Van der Panne, 2004; Johansson, 2005).

139 Henderson (2003a) finds that intra-sectoral specialization tends to have a positive effect on productivity.

140 See also Glaeser (2000), Feldman (2000) and DeGroot et al. (2009).

141 For complementary contributions that discuss this approach refer to Romer (1990b), Harhoff (1995), Feldman (1999), Keilbach (2000), Caniëls and Verspagen (2001), Neffke *et al.* (2009) and Neffke *et al.* (2011).

It is argued in several studies that concentrated industries are in most cases mature industries and technology fields that are organized in large scale production. MAR externalities are predominantly present in such mature industries (Feldman, 2000; Audretsch and Feldman, 2004).[142] Accordingly, mature industries are considered to have enjoyed from cross-fertilization and high growth rates of inventions at early stages of the life cycle (Audretsch *et al.*, 2008).[143] Nevertheless, an important shortcoming of the MAR-spillover concept is that it predominantly centers the origin of the externalities in an industrial perspective, i.e., intra-industry effects. However, the working channels (carriers) of such knowledge spillovers remain a "black box" and object for critique and further research (Feldman, 1999, 2000; Breschi and Lissoni, 2001a).

2.1.6.3. Jacobs Externalities and Diversity in Cities

"Jacobs externalities" are treated as a particular type of urbanization externalities.[144] They represent dynamic inter-industry effects that originate from a significant diversified production structure and are in most cases linked to inter-industry knowledge spillovers (Partridge and Rickman, 1999; Johansson, 2005; de Groot *et al.*, 2009).[145] Related to innovative capacity and innovation output, they are also labeled "innovation externalities" (Johansson and Quigley, 2003; Capello, 2007, 193). Accordingly, innovation externalities are dynamic economies which appear in the system as new routines (process innovations) or new products and an increasing product diversity (product innovation) (Johansson, 2005).[146]

In the late 1960s, Jacobs (1969, 71) suggested that

> "[o]ur remote ancestors did not expand their economies by simply doing more of what they already been doing [...]. They expand their economies by adding new kind of work. So do we. Innovating economies expand and develop. Economies that do not add new kinds of goods and services, but continue only to repeat old work, do not expand much nor do they, by definition, develop."

Following her arguments, cities have to frequently upgrade their industries (see also Audretsch and Feldman, 1999; Duranton and Puga, 2001).[147] Otherwise, a monotonous urban industry structure is considered to induce a stagnant settlement. As she has argued,

---

142 See also van der Panne and van Beers (2006), de Groot *et al.* (2009) and Neffke *et al.* (2011).

143 Griliches (1979, 104) presented the computer industry as an example: *"[The computer industry] has had a tremendous real productivity growth, most of it unmeasured in its official indices, and most of it unappropriated within the industry itself [...] because of rather intensive competitive pressures."*

144 See Taylor (2006) for an overview and appreciation of Jane Jacobs.

145 See also Combes (2000a), Keilbach (2000), Carlino *et al.* (2001), Henderson (2003a), Autant-Bernard and Massard (2007), Glaeser and Resseger (2009) and Dauth (2010).

146 This idea of cross-fertilization of industries has already been mentioned by Mansfield, who has linked the spillover approach to knowledge as a relevant input. He has concluded that *"[...] techniques invented for one industry turning out to be useful for others as well."* (Mansfield, 1968, 19). Similar ideas of inter-industry technology flows have already been argued by Scherer (1982).

147 See also Behrens and Robert-Nicoud (2008) and Henderson (2010).

> "[a] very successful growth industry poses a crisis for a city. Everything – all other development work, all other processes of city growth, the fertile and creative inefficiency of the growth industry's suppliers, the opportunities of able workers to break away, the inefficient but creative use of capital – can be sacrificed to the exigencies of the growth industry, which turns the city into a company town. [...] Monopolies gratuitously harm cities and suppress what their economies are capable of achieving [...] extortionate prices, harmful though they most certainly are, are the least of disadvantages of monopolies, for monopolies forestall alternate methods, products and services" (Jacobs 1969, 124-125).

As firms' (and regions') innovative capacities are assumed to be stimulated by inter-industry spillovers, the idea of cross-fertilization has been related to the regional settlement structure, which means that the industry dimension of agglomeration economies is combined with a geographic dimension (Audretsch and Feldman, 1999; Carlino, 2001; Capello, 2009). Large urban regions are considered to be more efficient and to gain in efficiency (productivity), to innovate faster (innovation rate) and to grow faster (employment growth) (Johansson and Quigley, 2003; Johansson, 2005; Capello, 2007, 195). Thus, cities and urban areas are regarded to raise innovative capacities because they act as arenas for the confluence of innovative factors (Fujita and Thisse, 1996; Duranton and Puga, 2001; Johansson and Quigley, 2003). According to the Schumpeterian paradigm, new economically relevant knowledge emerges from creating new combinations of existing ideas. Related to this view, Glaeser (2000, 83) considered cities to be centers of excellence for the creation and transmission of ideas and figured that

> "[c]ities will grow when they are producing new ideas or when their role as intellectual centers is increasing."

Glaeser *et al.* (1992) investigated how industrial diversity of a city - and not solely its size and the size of its industries – can give rise to growth promoting agglomeration economies.[148] The authors argued that, in contrast to static externalities, dynamic externalities have implications for industry growth rates in cities.[149]

Similarly, Audretsch and Feldman (1999) and Audretsch *et al.* (2008) linked higher potentialities for invention and innovation to urban, diversified industry structures but not to specialized ones. Their conclusions are based upon the empirical result that the number of new US product announcements of diversified spatial units exceeds those of industries located in cities, which are specialized into a few industrial activities.[150]

The Jacobs-externality approach is also in line with Florida's "creative class" hypothesis, where creative people are considered to locate in diversified locations and creative places

[148] In a later work, Glaeser (1996, 230) has pointed out the idea that *"[g]rowth hinges on the movement of ideas, naturally led to a re-exploration of the economic role of cities in furthering intellectual flows."*

[149] Glaeser *et al.* (1992, 1128) additionally argued that *"[dynamic externalities] are different from the more standard location and urbanization externality theories that address the formation and specialization of cities (Henderson 1986) but not city growth."*

[150] In a similar context Kelly and Hageman (1999) stated that the location of industry R&D activities is much more determined by other industries' R&D activities than industries' own production.

to (re-) combine pieces of knowledge and to establish an entrepreneurial society (Florida, 1995, 2002c,a).[151]

More recently, based on the above presented concepts, classifications and debates, researchers in a geography of innovation tradition are challenging the "relatedness" of industries (Boschma and Frenken, 2009b; Boschma and Iammarino, 2009; Neffke *et al.*, 2011). These studies test the hypothesis that knowledge predominantly spills over between related technology fields (i.e., related variety), which adds a cognitive dimension to the industrial and spatial dimension of agglomeration economies.[152]

To conclude, Jacobs externalities are generally linked to the benefits from industrial cross-fertilization and are in most cases associated with distance decay effects (proximity externality). Accordingly, it is argued that cities serve as central places for upgrading a region's innovative capacity because knowledge is assumed to spill over between industries at a proximate distance. Hence, accumulated knowledge of a specific industry can (partially) be applied in other (or related) industries.[153]

With a glance on the technological diversification and presence of different technology-specific research clusters in European capital and metro regions and urban and rural areas, this study applies EPO patent applications at the regional level to contribute to the "specialization-diversity" debate (see chapter 3).

2.1.6.4. Porter Externalities and the Competitive Advantage of Regions

In comparison to the inter- and intra-industry knowledge spillover debate, economists (and geographers) have discussed the effects of competition on the rate of innovation and growth; some argue that monopoly structures encourage innovation; others argue that competitive markets show higher rates of innovation (Carlino, 2001; Carlino *et al.*, 2001; Capello, 2007).[154] Porter argued in several studies, in line with the MAR-externalities concept, that knowledge spillovers and strong competition in geographically concentrated and specialized industries stimulate growth (Glaeser *et al.*, 1992). In this context, externalities are associated with cities and urban areas and originate from competition between proximate firms. The approach is based upon Porter's "competitive advantage concept" and the "Diamond approach," which mainly consists of four determinants (Porter, 1990, 1998a,b).[155]

According to Porter (1990, 71), the approach centers

[151] In this respect, the creative class approach resembles a sort of human capital theory.

[152] The "diversity-specialization-debate" is addressed in the empirical literature review (section 2.2) and challenged in the empirical analyses with special focus on research clustering and patenting activity (sections 3 and 3.5).

[153] For further details refer to Jacobs (1969), Glaeser et al. (1992), Audretsch and Feldman (1999), Breschi and Lissoni (2001), Scott and Storper (2003), Athreye and Werker (2004), Greunz (2005).

[154] According to Carlino (2001, 19), *"[w]hen local economies are competitive, the innovations of local firms are rapidly adopted and improved by neighboring firms. In contrast, local monopolists tend to rest on their laurels rather than risk innovation."*

[155] For a critical discussion refer to Glaeser *et al.* (1992), Maggioni (2002), Martin and Sunley (2003), Martin and Sunley (2005).

> "(i) factor conditions (e.g., the nation's position in the factors of production, such as skilled labour or infrastructure, necessary to compete in a given industry; (ii) demand conditions (e.g., the nature of home demand for the industry's product or service; (iii) related and supporting industries (e.g., the presence or absence in the nation of supplier industries and related industries that are internationally competitive); (iv) firm strategy, structure and rivalry (e.g., the conditions in the nation governing how companies are created, organised, and managed, and the nature of domestic rivalry."

Porter (1996, 87) adapted the Diamond concept, which has been introduced at the level of countries, to the regional level and suggested that

> "[r]egional clusters grow because of several factors: concentration of highly specialized knowledge, inputs and institutions; the motivational benefits of local competition; and often the presence of sophisticated local demand for a product or a service."

In rethinking the effect of location on competition, he argued that local competition is mostly limited to competition for natural resources, employees and inputs, e.g., highly-skilled employees (Porter, 2000). But effects from local competition on the international competitiveness of firms is hard to theorize (Martin and Sunley, 2003). Similarly, Glaeser *et al.* (1992) argued that fierce local competition may produce significant incentives to innovate faster. According to this, Porter externalities are very similar to the MAR-case presented above; the discussions, however, should focus on the underlying market structure.[156]

To summarize, Porter externalities are intra-industry externalities. Opposed to monopoly power, the presence of pure competition and rivalry in a cluster (or region) propels creation, adoption and diffusion of information and knowledge and the development of new products. According to this approach, the incentives to innovate are greatest when markets work under strong competition.[157] Unfortunately, competition and market structures cannot be explored in the pan-European context in this study due to a significant lack of firm-level data.

2.1.6.5. A Taxonomy of Innovation Externalities

The above presented different sources and working channels of technological externalities are summarized in table 2.4. The main focus is on knowledge spillovers and information externalities that originate (i) from anonymous market-led activities as an unintended by-product, (ii) from routinized and sustained transaction linkages between well-known partners at a proximate or long-distance (i.e., R&D co-operations, co-patenting networks), or (iii) from MAR- and Jacobs externalities (knowledge spillovers) in agglomerations (proximity externality). With respect to the latter, a relevant source of knowledge externalities is related to unintended spillovers from knowledge providers in dense agglomerations.

---

156 Glaeser *et al.* (1992, 1127-1128) argued that *"[k]nowledge spillovers in specialized, geographically concentrated industries stimulate growth. (Porter) insists, however, that local competition, as opposed to local monopoly, fosters the pursuit and rapid adoption of innovation."*

157 For a discussion see also Audretsch and Feldman (1999), Audretsch and Feldman (2004) and Feldman and Kogler (2010).

The MAR- and Jacobs externality approach emerged in the mid 1980s and 1990s and reflects a combined industrial and geographic dimension on (dynamic) agglomeration economies that primarily differentiates between inter- and intra-industry effects in a spatial context (Audretsch and Feldman, 1999; Capello, 2009). However, the main focus of the approach is on scale and diversity but not explicitly on the working channels and micro-foundations of knowledge transmission (Breschi and Lissoni, 2001a). Although it is a multidimensional approach that includes the industry and geographic dimension, it gives no role to synergies, to research networks, to innovation and learning or to socio-cultural or cognitive aspects.

Nevertheless, inter- and intra-industry innovation externalities may also originate from tacit knowledge transmission in localized social (informal) networks and from long-distance research collaborations, which is in line with the "collective learning approach" (McCann *et al.*, 2002; Johansson, 2005; Wilhelmsson, 2009).[158] Therefore, section 2.1.7 places the emphasis on these shortcomings.

2.1.6.6. Endogenous Growth Theory and Research Clustering

2.1.6.6.1. Knowledge Stocks and Knowledge Spillovers

In order to explain persistent national and regional growth disparities and the emergence and persistence of research clustering, the relationship between research activities, knowledge stocks and innovation processes are taken into account in the literature. Early endogenous growth (NGT) models are mainly built upon technological externalities and thus correspond to the proposed taxonomy of agglomeration economies and externalities (see sections 2.1.4.3 and 2.1.6).[159]

An influential contribution to the endogenous growth literature represents the work of Paul Romer (1986, 1987, 1990b,a). Romer himself explained the idea of regional disparities as rooted in a Smith-Marshall-Young-Kaldor tradition (Rima, 2004; Chandra and Sandilands, 2005; Solow, 2007).[160] What has been essential for the cluster literature, in the context of knowledge-intensive industries, is the combination of ideas about spatial interaction with theoretical aspects of endogenous growth (Capello, 2007; Eckey, 2008; OECD, 2009a). Following the early contributions of Romer (1986), ideas (or knowledge) generally do not correspond to the law of diminishing returns as opposed to labor or capital inputs (see section 2.1.7). In Romer's original NGT version (1986), the stock of knowledge in the economy is intertwined with the capital accumulation process, but the "public" capital

[158] The diffusion of knowledge in social networks is also addressed in section 2.1.7 and the empirical review in section 2.2.6. Refer to Johansson and Quigley (2003), de Groot *et al.* (2009), Breschi and Lissoni (2009), Capello (2009).

[159] For a comprehensive overview refer to Capello (2007), Eckey (2008) and OECD (2009a).

[160] Romer (1986, 1004) argued that *"[t]he idea that increasing returns are central to the explanation of long-run growth is at least as old as Adam Smith's story of the pin factory. With the introduction by Alfred Marshall of the distinction between internal and external economies, it appeared that this explanation could be given a consistent, competitive equilibrium interpretation. The most prominent such attempt was made by Allyn Young in his 1928 presidential address."*

remains uncompensated. That being the case, the technological progress and interdependencies between new ideas, technological knowledge and capital accumulation induce endogenous growth (Romer, 1986; Capello, 2007).

Table 2.4. Innovation externalities

Externality type	**Transaction type**	**Mechanism**
Industry structure in agglomerations	Spillovers based on proximity; different innovation spillovers in an agglomeration; MAR- and Jacobs externality (see sections 2.1.6.2, 2.1.6.3 and 2.1.7.3).	Size and structure of the agglomeration; inter- and intra-industry knowledge externalities stimulates the innovation processes and product development of firms.
Market vs. networks (section 2.1.7.7)	Upstream knowledge spillovers. Unintentional spillovers as by-product of anonymous market transaction.	Information and/or knowledge spills over as by-product of interaction between an (knowledge) input-buying firm and its suppliers.
	Downstream knowledge spillovers. Unintentional spillovers as by-product of anonymous market transaction.	Information and knowledge spillovers occur as a by-product of interactions between an (knowledge) input-selling firm and its customer.
	Upstream knowledge spillovers. Routinized, persistent inter-firm (network) transaction linkages induce knowledge externalities.	Information and/or knowledge spillovers from a vertical transaction linkages (formal network); directed spillovers stimulate innovation output of the buyer.
	Downstream knowledge spillovers. Routinized, persistent inter-firm (network) transaction linkages induce knowledge externalities.	Information and knowledge spillovers from a vertical transaction linkages (formal network); directed spillovers stimulate innovation output of the seller.
Competition	Competition between proximate companies in a specialized cluster (Porter externality, see section 2.1.6.4).	Due to physical proximity competing firms imitate competitors in order to move towards best-practice, improve routines and develop new products.

Source: illustration based on Johansson (2005, 122), Johansson and Quigley (2003) and Capello (2007, 2009).

Knowledge is assumed to enter the production sphere in two ways (Romer, 1986; Capello, 2007, 242). First and foremost, newly developed technological knowledge is utilized by the firm that has invested in its development to obtain productivity effects and new products.

This knowledge can be protected from being imitated. Second, the new pieces of knowledge that have been protected by the legal system increase the stock of public available knowledge as the technological knowledge is codified in the patent application. Thus, designs become available for other researchers in the form of public patent documentations (i.e., blueprints). By studying the blueprint, pieces of knowledge have a high propensity to spill over to other researchers (and regions/clusters) and will increase their productivity. Moreover, one could think of cross-fertilization of firms in other industries (see sections

2.1.4.3 and 2.1.6). The "diversification-specialization debate" exactly builds upon this argument (Beaudry and Schiffauerova, 2009; de Groot *et al.*, 2009). The question then centers the issue of the spatial and technological proximity of knowledge (i.e., technological relatedness). Several authors discuss the threshold level of spatial and technological (and cognitive) distance (Boschma and Frenken, 2009b; Neffke *et al.*, 2009, 2011).[161]

Regarding the main mechanisms, most contributions to the endogenous growth theory build upon the concept of technological externalities, which is specified by non-rivalry and (partial) non-excludability, e.g., in a blueprint-producing R&D sector (Arrow, 1962b; Romer, 1990b; Jones, 2004).[162] Introducing externalities converts decreasing returns into constant or increasing ones.

In a regional context, regions with high population densities and a crucial size of their populations of researchers are expected to reach high levels of innovative output (innovation externality) and efficiency via new routines (efficiency externality). Moreover, it is assumed that researchers generally build dense and developed social networks that increase the scope for the exchange of information, ideas and technological knowledge (Breschi and Lissoni, 2009). These prerequisites could then stimulate knowledge exchange and technological progress in the spatial unit (Gordon and McCann, 2000; Iammarino and McCann, 2006).[163] These thoughts are reflected in the "proximity versus network debate," which focuses on innovation output and productivity growth (Johansson, 2005; Capello, 2007).

2.1.6.6.2. Technological Externalities and Specialization

Several endogenous growth models interpret the source of increasing returns as rooted in fixed costs and/or technological externalities.[164] Accordingly, besides the pure knowledge spillover approach (Romer, 1986), increasing returns are linked to the concept of fixed costs, which implements a form of indivisibilities (non-convexities) and specialization (Romer, 1987, 1990b,a, 1991; Rima, 2004; Chandra and Sandilands, 2005).[165] Therefore, these models combine the mechanisms presented in sections 2.1.4.3, 2.1.5 and 2.1.6.

---

161 As has been argued by Castellacci (2008, 985), *"[t]he general proposition that innovation and inter-sectoral knowledge spillovers are important for the international competitiveness of manufacturing industries is a major point of agreement between new growth theories and evolutionary economics. The two approaches, however, differ substantially in terms of the conceptualization of the innovative process and the analysis of its economic impacts."*

162 Regional knowledge bases can be non-rival because they can be utilized by agents without limiting their use by additional agents (non-rivalry). Moreover, non-rivalry goes in most cases hand-in-hand with non-excludability (Romer, 1990b). This circumstance distinguishes knowledge and information from capital goods and equipment, which can normally only be used in one location (at a certain time).

163 This view will be challenged empirically in sections 3.5 and 4.3.5.

164 For a detailed review of the NGT refer to Seiter (1997).

165 As mentioned by Romer, the concept of Young's increasing returns is rather built upon the division of labor; this concept combines tendencies of specialization and the expansion of the market, opposed to increasing returns by assuming fixed costs due to economies of scale. Romer (1989, 198) has argued that *"[t]he degree of specialisation, or equivalently, the number of different firms that are available at any point in time or location, is limited by the presence of fixed costs [...] Although Marshall and Young choose to describe specialization in terms of competitive equilibrium, with externalities, it is now clear that a more rigorous way to capture the effects they had in mind is in a model with fixed costs. In an equilibrium with nonnegative profits, price must exceed marginal costs to be able*

Based on his early ideas, Romer (1990b) has contributed with a model of endogenous growth, which includes fixed costs (non-convexities), knowledge spillovers and a monopolistic market structure. The propelling drivers in the model are (i) horizontal product innovations, (ii) monopolistic competition and (iii) R&D activities and sector-wide learning effects on researchers' productivity (see also Rima, 2004; Chandra and Sandilands, 2005).[166] The framework does not explicitly contain capital goods but builds upon high-skilled labor (human capital) and blueprints in the production process. Thus, knowledge is assumed to enter the system via human capital (rivalry) and blueprints (non-rivalry). The model represents a three-sectoral system. It consists of a perfectly competitive final good sector (manufacturing good) that uses intermediates; the intermediate sector uses the new designs.[167] To implement a sort of learning process, productivity in the R&D sector is assumed to grow proportionally with the accumulated number of blueprints, i.e., the stock of designs, which leads to positive technological externalities. Knowledge spillovers are thus assumed to drive productivity gains in the R&D sector (sector-wide learning curve).[168] This is the so-called "standing on shoulders effect," which enables endogenous growth in these models (see also section 2.1.6.7). The cumulative process shows some kind of scale effect. The higher the knowledge stock of a region, respectively the stock of blueprints in a region, the higher are the associated productivity gains in the region-specific research sector. If the stock of knowledge is a function of the number of researchers in a region, then agglomerated regions (cities, metropolises) should show higher rates of innovation (Eckey, 2008, 132).[169] The number of firms (which is equal to the number of individuals) is determined by the regional population; the number of firms, entry rates and the scale of operation is thus not endogenous. There is no entry because labor supply (i.e., the number of researchers) is fixed (Eckey, 2008).[170]

2.1.6.6.3. Conclusions and Critical Remarks

Similar to the above described model, Grossman and Helpman (1991b, 1993) have shown that localized spillovers can lead to geographic disparities and research clustering. The assumption of (partial) non-excludability of knowledge suggests that R&D activities can induce technological externalities (i.e., knowledge spillovers and their productivity effects)

to recover these fixed costs, so the model must therefore contemplate some forms of market power." For comprehensive overviews refer to Roos (2002b), Rima (2004), Chandra and Sandilands (2005), Capello (2007), Eckey (2008), Harris (2008) and OECD (2009a).

166 It is, however, essential to understand that the model of Romer (1990b) includes no machinery sector or capital goods in the production function. Most capital theories introduce technological progress via capital good embodied technological change by means of increasing productivity.

167 The assumption of an R&D sector is picked up in several path-breaking models in new economic geography and endogenous growth theory. Finally, the NEGG model in this paper also has an R&D sector that produces designs.

168 Romer (1986) also used externalities; however, his initial paper of 1986 does not include monopolistic competition and deals with an aggregated production function and global knowledge spillovers.

169 Consequently, several models showed scale effects, where an increasing population of entrepreneurs or higher population growth rates would lead to higher productivity gains/regional growth rates.

170 The simultaneous treatment of agglomeration and growth is not the primary concern. For this purpose, agglomeration and growth need to be combined via the introduction of knowledge and its diffusion in space. Industrial specialization, clustering and spatial diversity then highly depend on knowledge exchange and knowledge externalities.

(see section 2.1.7). However, the primary interest of the authors is grounded on explaining endogenous growth but not explicitly agglomeration economies and research clustering (and centripetal and centrifugal forces).[171] Grossman and Helpman (1993, 16) argued that

> "[b]y technological spillovers we mean that (1) firms can acquire information created by others without paying for that information in a market transaction, and (2) the creators or current owners of the information have no effective recourse, under prevailing laws, if other firms utilize information so acquired. [...] The technological spillover that result from commercial research may add to a pool of public knowledge, thereby lowering the cost to later generations of achieving a technological breakthrough of some given magnitude. Such cost reductions can offset any tendency for the private returns to invention to fall as a result of increases in the number of competing technologies."

Endogenous growth models have been heavily criticized. Jones (1999) argued that regional growth in the endogenous growth theory is problematic as it is driven by the implementation of externalities that originate from the stock of accumulated ideas (blueprints) and discoveries due to non-rivalry (see also Keilbach, 2000; Jones, 2004).[172] The stock of ideas is assumed to be directly proportional to the economy's research effort, which, in turn, is considered to be a function of the total population, i.e., researchers or the creative class (Florida, 1995, 2002c; Jones, 2004; Glaeser, 2005a). Therefore, a considerable number of endogenous growth models are accused to suffer from the so-called "scale effect" (Jones, 1999). It is the link between ideas and returns to scale that gives rise to a basic scale effect in idea-based growth models (Romer, 1986, 1990b; Grossman and Helpman, 1991b; Aghion and Howitt, 1992). The growth rate of blueprints (and the economy) is proportional to the amount of research activities undertaken in the economy which is itself dependent on the growth rate of the population. Accordingly, it is argued that an increase in the size of the regional population, other things being equal, raises the number of researchers and therefore increases the growth rate of per capita income.[173]

Unfortunately, the transfer mechanisms of spillovers remain a "black box." The models generally abstract from the manifold transfer channels as will be highlighted in section 2.1.7. Much of the criticism has also centered on the non-excludability assumption (Capello, 2007; Freund, 2008). In many cases, economically useful technological knowledge fulfills the attribute of partial excludability, because it is possible to prevent other agents from using technologies (e.g., epistemic communities). Excludability is reflected by technological patterns (tacit knowledge, secrecy) but also by the legal system in an economy. Accordingly, knowledge can be made partially excludable by patent systems. However, depending on the technology case, some pieces of knowledge always remain inaccessible because codification is complex and some additional parts of knowledge have a tacit nature (section 2.1.7.2).

In conclusion, although the economization of knowledge is protected by legal systems (patent systems), spillovers of knowledge may lead to productivity effects in the develop-

---

171 See also Baldwin and Martin (2003), 19, footnote.

172 Romer (1990b, 98) argued in his 1990 paper that *"[the model here suggests that what is important for growth is integration not into an economy with a large number of people but rather into one with a large amount of human capital."*

173 Jones (1998, 11) argued that *"[t]he size of the economy affects either the long-run growth rate or the long-run level of per capita income."*

ment of further products and processes (section 2.1.6.7).[174] Further to this, if the potential output of the researchers, who are studying blueprints in the patent documents of their rivals, could violate existing intellectual property rights in the future is rather a technical and legal issue (i.e., the IPR debate). After all, it is clear that patent documents contain codified knowledge (see section 2.1.7), which is available for those who can translate and use it, thus narrowing knowledge transmission to a technology- and sector-specific process and flows within epistemic communities. The translation of blueprints imposes considerable costs and presupposes specific knowledge and skills (partial excludability) only existent in epistemic communities (Steinmueller, 2000; Cowan *et al.*, 2000; Lissoni, 2001). This assumption is consistent with the "absorptive capacity concept" (Cohen and Levinthal, 1990). Related to these thoughts, scientific/technological knowledge is considered to be rather a club good as only fractions of researchers (and firms) can access, translate, modify and recombine such pieces of knowledge (Freund, 2008). Growth poles and centers of technological excellence will persist if the local stock of technological knowledge, which is strongly connected to the number of researchers, is unequally distributed across space. Moreover, the clustering of research activities may persist under the assumption of spatially localized knowledge externalities (proximity externalities), which is strongly related to some key properties of innovative milieus (section 2.1.7.3). An initial non-symmetric distribution of knowledge stocks (i.e., blueprints, researchers, research laboratories) and/or strong distance decay effects of knowledge transfer (i.e., pure externalities and partially compensated flows) may be able to explain the persistence of core-periphery structures, research clustering and regional growth differentials (Capello, 2007; Sachs and McCord, 2008; Henderson, 2010). According to these thoughts, endogenous growth models can be used to explain income and growth differentials and persisting regional disparities in research activities, but they are not suited to explaining the emergence of regional core-periphery structures, at least without migratory movements or researchers, changing distance decay effects of knowledge transmission or other centripetal forces that induce cumulative circular causation.[175]

The spatial distribution of knowledge stocks and researchers is a crucial factor for regional development. Therefore, the distribution (and clustering) of knowledge stocks and disparities in research and patenting activity in Europe will be explored in this study by analyzing the concentration of EPO patenting activity at the regional level (see chapter 3). Persistent core-periphery structures in patenting activities should then be reflected in significant differences regarding regional growth rates (see chapter 5).

2.1.6.7. Research Clustering and Knowledge Flows in Core-Periphery Models

2.1.6.7.1. Agglomerations, Blueprints and Technological Externalities

In the standard NEG framework, second-nature causes of agglomeration and clustering solely capture pecuniary causes and effects of agglomeration, e.g., vertical linkages, in-

[174] The literature then issues a "free lunch" from positive technological externalities.
[175] The role of innovative capacities, knowledge stocks and research clustering with respect to GDP growth rates of European regions will be discussed in chapter 5.

creasing returns, transport costs and distance sensitive production, and mobility of workers.[176] Recent contributions focus on the concentration of production and R&D activities, which induce alternative centripetal forces (i.e., R&D spillovers) and determine relocation and regional growth.[177] Clustering of research activity can be traced back to several factors, e.g., accesses to (i) market information, (ii) codified and tacit knowledge, (iii) skilled employees (and human capital in general), (iv) patenting and researcher networks (see section 2.1.7.5), and (v) specialized suppliers in highly fragmented production and supply chains. Accordingly, research activities show indeed spatial concentration in several industries, since firms and entrepreneurs are inclined to co-locate in those regions where required inputs, tasks and processes co-agglomerate (Feldman, 1994b; Gallagher, 2008). The discussed spatial nature of knowledge (see section 2.1.7) has directly entered the new economic geography literature and established a second generation of models by means of an explicit recognition of knowledge transmission as a pivotal centripetal force.[178] Fujita and Krugman (2003, 161) concluded that

> "[t]here recently appeared several multiregional growth models such as Martin and Ottaviano (1999), Baldwin et al. (2001) and Fujita and Thisse (2003) in which a core-periphery model is grafted onto a Grossman-Helpman-Romer-type model of endogenous growth. [...] the proximity of people is certainly helpful in the diffusion and generation of knowledge (in particular, through face-to-face communications)."

In this respect, Baldwin and Martin (2004) and Gosens and de Vaal (2010), among others, point out the importance of face-to-face contacts for knowledge transmission (section 2.1.7.2) and localized knowledge flows (section 2.1.7).[179] Accordingly, agglomeration of manufacturing industries, R&D activity and regional growth in blueprints (i.e., capital goods) seem to be positively related. That being the case, agglomeration is considered to have a positive effect on regional growth (i.e., growth in blueprints). Knowledge stocks and manufacturing industries co-locate in regions, which induces research clustering.[180] As a consequence, regional disparities remain and non-equity of research distribution is a possible and maybe stable outcome.[181] The centripetal force which originates from R&D concentration and localized knowledge spillovers may be quite strong. Given the assumption of constant (non-decreasing) returns to learning in the aggregate production function of regions, knowledge spillovers in agglomerations can be interpreted as an elementary source of sustained regional growth (Glaeser *et al.*, 1992; Fujita and Thisse, 1996, 2003).

176 Pecuniary externalities are transferred and stimulated via the market mechanism. However, they are different from technological externalities as the latter explicitly focus on non-rivalry and non-excludability of knowledge as an input.

177 Baldwin and Martin (2003, 28) suggested that *"[...] growth affects geography which itself affects growth and agglomeration is driven by the appearance of growth poles and sinks."*

178 According to the OECD, *"[k]nowledge can be regarded as an economic output in the form of a production blueprint but knowledge is also an input required to produce new blueprints. [...] In this sense, it recalls a corn-economy in which corn produces corn"* (OECD, 2000, 21).

179 They refer to the existence of some kind of "home-bias."

180 In other words, *"[g]rowth, through innovation, spurs spatial agglomeration of economic activities which in turn leads to a lower cost of innovation and higher growth so that a circular causation between growth and the geographic concentration of economic activities sets in"* (Martin and Ottaviano, 2001, 948).

181 These results are in line with the findings from the spatial variants of endogenous growth models developed by Romer (1990b,a) and Grossman and Helpman (1991a), among others.

Furthermore, the regional disparities of knowledge spillover effects between the center and the periphery explain diverging growth paths in blueprint production and the intensity of R&D clustering. According to this conceptualization, the implementation of knowledge externalities and sectoral learning processes can be regarded as the foundation of a second generation new economic geography framework, the "new economic geography growth" (NEGG) or "growth-cum-geography models" (Baldwin and Krugman, 2004).

However, the NEGG models show some central differences regarding centripetal and centrifugal forces. The subsequent table 2.5 summarizes the forces that influence the emergence and stability of core-periphery structures and the concentration of research activity.[182] In general, the new economic geography can be divided into three classes of models.[183] The seminal contribution was made by Krugman (1991) with his core-periphery framework that allows for labor mobility and expenditure relocation (see section 2.1.5.5). Later on, the core-periphery framework was enriched by vertical linkage (VL) models that focus on centripetal forces originating from input-output linkages, by footloose capital (FC)/capital accumulation (CC) models that center R&D externalities that induce cost linkages and by footloose entrepreneur (FE) models that place the emphasis on migratory movements of agents (Cerina and Pigliaru, 2005; Cerina and Mureddu, 2009).

In the following, selected models that challenge the regional distribution and structural dynamics of research and patenting activity are briefly presented and discussed.

2.1.6.7.2. Growth-Cum-Geography Models and R&D Location

Baldwin and Forslid (2000a) introduced a cum-growth-geography framework based on R&D spillovers. Spillovers can either be intra- or inter-regional (but also somewhere in between). Blueprint (capital) mobility between the regions and repatriation of profits is impossible. A high level of transportation costs (i.e., a low level of integration) is always a stable equilibrium with no core-periphery emergence; however, it is accompanied by lower regional growth rates. The authors showed that a change in transportation costs (at the break point) leads to a "growth take-off" stage if one region accumulates relatively more blueprints (i.e., knowledge). The regional growth rate increases until a new equilibrium is reached. At higher levels of transportation costs, the model predicts full industry agglomeration in only one region; accompanied by higher growth rates (i.e., growth in blueprints) but also increasing regional disparities between the two regions, which again induces normative issues and a trade-off between equity distribution of industrial activity and aggregate

182 It should be realized that the last row centers the spatial range of technological externalities or knowledge spillovers that are crucial in NEGG models. The other factors, however, also influence core-periphery stability.

183 The conceptualization of different geographical scales in models demonstrates that agglomerations are regarded to be influenced by varying centripetal and centrifugal forces. Thus, all these forces form and modify the spatial complex system of economic activity. The essential contribution of NEG models is then to devise a modeling approach that can give essential ideas and information about the centripetal forces that pull the economy together and the opposed centrifugal forces that push things apart (Fujita and Krugman, 2003). As Combes *et al.* (2005, 330) concluded, *"[...] there is no inherent contradiction between the urban system approach and NEG: the latter is trying to explain broad trends at large spatial scales while the former attempts to explain spikes of economic activity."*

Table 2.5. Cumulative causation and forces of agglomeration

Centrifugal Forces	**Centripetal Forces**
a small regional market and small initial manufacturing share	thick markets, home-market effect (i.e., large expenditure share) and expenditure shifting
immobile factors of production (i.e., inter-regional immobility of, e.g., labor, entrepreneurs, patents)	mobility of entrepreneurs, consumers, firms, patents/blueprints (factor mobility)
competition effect (falling price index in the agglomeration)	initial (exogenous) higher level (skewness) of factor endowments (first-nature)
land rents, commuting, congestion costs	intra-regional vertical linkages (intermediates, resources, skilled labor input)
global knowledge spillovers (public good, no distance decay); inter-regional social ties; inter-regional co-inventor linkages	local knowledge spillovers (strong distance decay effects); intra-regional social ties; intra-regional co-inventor linkages

Source: own illustration based on Baldwin *et al.* (2001b), Roos (2002b), Baldwin and Martin (2004), Cerina and Pigliaru (2005), Cerina and Mureddu (2009) and Gosens and de Vaal (2010).

growth. To conclude, the model shows that regional growth rates are affected by research and industry clustering.[184]

In Baldwin and Forslid (2000b), a Krugman-type geography framework is merged with an endogenous growth model. The monopolistically competitive sector (increasing returns) uses labor and capital inputs (i.e., blueprints). The production of blueprints exhibits technological externalities in terms of sector-wide (localized) learning effects; thus, industry agglomeration and research clustering are considered to be growth enhancing. Accordingly, localized knowledge spillovers and the growth of blueprints are considered to enforce the agglomeration process. Skilled workers (and their expenditures) show migratory movements between the regions due to varying present values of the underlying utility function, which induces expenditure shifting (centripetal force) (see also Martin and Ottaviano, 1999). To conclude, endogenous growth emerges from sector-wide learning effects in the R&D sector and core-periphery is additionally driven by migratory movements, which differentiates this model (and others) from early endogenous growth models. The model demonstrates that localized knowledge spillovers and research clustering generally support the emergence and stability of core-periphery structures (independent from the level of transportation costs), whereas global spillovers can also induce a process of dispersion. Regarding these predictions, it seems to be of great importance to measure research clustering in Europe.

Martin and Ottaviano (2001) introduced a core-periphery model in which labor is assumed to be inter-regionally immobile. Labor inputs are used to produce a homogeneous consump-

184 For an overview of NEGG models also refer to Roos (2002b) and Litzenberger (2007).

tion good, a differentiated good and blueprints.[185] The authors merged the NEG with a Rivera-Batiz and Romer (1991) endogenous growth model. As usual in these frameworks, aggregate growth and the costs of blueprints are considered to be dependent on the total number of past research activities within the R&D sector. Blueprints are protected infinitely by a patent whose initial property belongs to agents in the region where the research effort was identified. Patents can be sold and are initially equally distributed among regions. The cost of R&D will be lower in the region where more firms are located due to localized knowledge spillovers. Both regions will engage in research activity if the manufacturing industry is equally distributed. That being the case, there are no incentives to relocate production of the increasing returns sector because the demand for varieties as well as their profits are the same in both regions. If the distribution is not equal, R&D activity also shows a core-periphery structure. If research activity is concentrated in one region, firms will tend to relocate to the core (agglomeration) where the local expenditure level is higher. Accordingly, agglomeration is an increasing function of growth (i.e., "forward linkage"). Furthermore, higher concentration of industries in one region is regarded to affect the R&D learning curve and thus to reduce the costs for additional blueprints (R&D externalities), which attracts more researchers (and firm entries) until profits are zero. Growth is modeled as an increasing function of agglomeration (i.e., "backward linkage").[186] If the initial distribution is asymmetric, the only stable steady state is the one in which all research activity is concentrated in a single region. This implies that production and R&D activities are geographically concentrated, even if research is still more agglomerated than manufacturing activity (see, e.g., Audretsch and Feldman, 1996).

Fujita and Thisse (2003) introduced an alternative core-periphery framework, which is similar to the one introduced by Baldwin and Forslid (2000b). The authors reported very similar conclusions concerning the agglomeration-and-growth relationship. A Krugman-type geography framework was extended to incorporate endogenous growth via blueprint accumulation (see also previous models). The R&D sector makes use of skilled labor to create new blueprints; the blueprints are then used in the manufacturing sector as a capital input. Similar to Baldwin and Forslid (2000b), the migration behavior of skilled agents is taken into account. Fujita and Thisse showed that industry co-location (and research clustering) leads to higher growth (in varieties). Moreover, if several centripetal forces are strong enough, even those firms are better off which remain in the periphery, although absolute discrepancies between the core region and the periphery increase (i.e., in industrialization, employment structure, wages).[187]

Concerning the possibility of cross-fertilization and inter-industry productivity effects, Cerina and Mureddu (2009) modified the framework of Baldwin *et al.* (2001a) by allowing inter-sectoral technological externalities (see also section 2.1.6.3). Their framework allows for spillovers from the R&D sector to the service sector. As a result of these inter-sectoral knowledge spillovers, the effects and welfare implications are even more unclear, compared

[185] The homogeneous good is produced under constant return to scale and perfect competition and transportation costs are nil. The manufacturing good, in opposition, is produced under increasing returns to scale and monopolistic competition and transportation induces costs (Martin and Ottaviano, 2001).

[186] The symmetric equilibrium (manufacturing share being equal) is only stable for positive equilibrium growth rate, when the regions start in a symmetric distribution.

[187] Indeed, this resembles a strong welfare conclusion similar to the ones included in the former NEGG models.

to the standard case without cross-fertilization. Assuming an initial core-periphery structure with an agglomerated industry in one region, the authors presented different effects. The core region is generally better off and benefits from two dynamic gains from clustering: (i) an increase in the nominal growth rate of blueprints due to localized knowledge spillovers; (ii) a decrease in the costs of services due to (localized) inter-sectoral knowledge spillovers (i.e., productivity effects). At the same time, the deindustrializing region experiences two effects: (i) a dynamic gain given by the increase in the nominal growth rate of manufacturing goods, which are transported to the periphery; (ii) a dynamic loss due to the fact that the stock of capital goods (i.e., blueprints) in the periphery does not grow anymore. Moreover, service prices are fixed as services do not benefit from localized productivity spillovers because inter-sectoral spillovers are localized in the core region. The authors assumed that these losses in the periphery may be counterbalanced by the gains in the services in the core region. Thus, agglomeration is at least welfare enhancing at the aggregate level.[188]

A final consideration concerns the implementation of social ties and networks. Gosens and de Vaal (2010) have built on the Fujita and Thisse (2003) "growth-cum-geography model" to analyze the implications of migratory movements and inter- and intra-regional social ties. Inter-personal linkages are considered to be a crucial factor for tacit knowledge exchange within and between regions (see section 2.1.7.2). They introduced a relationship between migration, knowledge spillovers and the concept of codified and non-codified (tacit) knowledge.[189] Codified knowledge is accessible and applicable to other researchers in the community and thus spills over quite easily, whereas non-codified knowledge may remain tacit and local due to several reasons, e.g., for lack of absorptive capacity and/or social ties.[190] According to Gosens and de Vaal (2010), migration is a key factor to benefit from knowledge from other regions. Moreover, inter-regional and intra-regional social ties lower the costs of exchanging tacit knowledge and ease knowledge spillovers, meaning that the region of immigration essentially profits from tacit knowledge of immigrant researchers when social ties to arriving researchers grow considerably.[191] The authors concluded that agglomeration of high-skilled researchers and thus research clustering in one region is not a straightforward outcome as the region of immigration and the region of emigration are both affected by existing knowledge networks. Accordingly, if immigrants continue their social relationships with the region of emigration, their newly accumulated tacit knowledge will spill over from the region of immigration to the region of emigration. This idea is identical to the arguments discussed in section 2.1.7.4.[192]

188 Considering the deindustrializing region, the real growth rate in case of the symmetric equilibrium can be higher compared with the core-periphery equilibrium if the dynamic loss in the services sector overcomes the dynamic gain in manufacturing. However, the standard NEGG interpretations find that the core and the periphery enjoy the same dynamical gains of agglomeration.

189 For a critical survey refer to Lissoni (2001) and Breschi and Lissoni (2001b).

190 According to the authors, tacit knowledge exchange is related to social interaction within the labor force, while productivity effects from codified knowledge is related to the number of manufacturing varieties in the region.

191 Important empirical studies are, e.g., Agrawal *et al.* (2006) and Saxenian (2006).

192 This reasoning resembles Saxenian's "new argonauts" that propel knowledge exchange between the United States and Asian countries (Saxenian, 2006).

2.1.6.7.3. Critical Remarks and Discussion

NEGG models can be applied to work out general mechanisms that induce relocation and geographic clustering of knowledge bases, researchers and patenting activity (i.e., research clustering). These frameworks have strong similarities to the geography of innovation literature (Audretsch and Feldman, 2004; Boschma and Frenken, 2009b; TerWal and Boschma, 2009) and the regional systems of innovation (RIS) approach (Cooke *et al.*, 1997; Cooke, 2008), although the methodological realization is different. All these research lines generally address clustering and challenge the observed spatial structures of innovative activity, research clustering and regional disparities relating to knowledge-intensive activity.

The aforementioned core-periphery frameworks offer important insights into the distributional dynamics of research clustering. Opposed to early NEG models (see section 2.1.5.5), recent contributions are additionally emphasizing the relationship between research clustering strength, agglomeration and growth of the R&D sector. The modeling structures, as presented in table 2.5, are different to early NEG models. Demand- and supply-side linkages, which emerge from pecuniary externalities in early NEG models, are replaced and/or complemented by alternative core-periphery mechanisms, i.e., non-pecuniary (technological) externalities.[193] In line with endogenous growth models, the details on knowledge transfer mechanisms remain a “black box.” Nevertheless, the main conclusion of NEGG models is quite similar to endogenous growth models: localized knowledge spillovers (and flows), which induce productivity effects in the R&D sector, sustain core-periphery structures. Therefore, it can be concluded that the spatial distribution of research activity, the distance decay of knowledge spillovers and the structure of inter-regional research networks, i.e., patenting networks, play a crucial role for regional development as major fractions of knowledge are distributed via (R&D-) network linkages. Unfortunately, network linkages still represent a minor framework alternative and an ignored channel of knowledge diffusion.

As has been emphasized by recent studies, inter- and intra-regional research networks may be well reflected by researchers’ co-patenting activities (see section 2.2). Regarding theoretical considerations of research networks, network mechanisms and knowledge flows via linkages represent a meaningful shortcoming of the above described models. The following section briefly summarizes major theoretical aspects of inter-personal and inter- and intra-regional research collaboration linkages and networks.

To conclude, from a empirical point of view, there is a great need for analyzing the distributional dynamics of European research activities at the regional level and to explore the structural dynamics of co-patenting network linkages between European regions. Therefore, the analysis of the distribution of research clusters and inter-regional co-patenting networks in chapters 3 and 4 is regarded as to challenge the aforementioned shortcomings in a European context.

193 The main sources and references to the NEGG model are Baldwin and Forslid (1999), Martin and Ottaviano (1999), Baldwin and Forslid (2000b), Baldwin *et al.* (2001b), Baldwin *et al.* (2001a), Baldwin and Martin (2003) and Cerina and Pigliaru (2005).

2.1.7. Agglomerations, Networks and Knowledge Transmission

2.1.7.1. Knowledge Flows, Network Linkages and Spillovers

Besides the broader concept of agglomeration economies presented in the previous sections (sections 2.1.4.3, 2.1.5 and 2.1.6), the survey is now related to the specific factors that are considered to represent the pivotal determinants of knowledge transmission, research clustering and the distribution of research activities. An important consideration concerns the theoretical conceptualization of the working channels of knowledge transmission in the context of spatial and relational proximity in milieus and the exchange of knowledge via inter-regional (trans-territorial) networks (Breschi and Lissoni, 2001b,a; Capello and Faggian, 2005). A crucial aspect in this regard consists of the attributes and properties of knowledge (Lissoni, 2001; Capello, 2009).

From a conceptual point of view, Tödtling *et al.* (2010), among others, have proposed a comprehensive classification of knowledge-sourcing strategies that differentiates between the possible knowledge acquisition channels at the firm level (see table 2.6). Similarly, Johansson and Quigley (2003) and Johansson (2005, 133) differentiated between agglomeration economies and the external economies of networks, whereby agglomeration economies and network economies are different but sometimes overlapping. They concluded that spillovers, which originate from networks and transaction linkages (quasi-market), are different from unintentional spillovers that emerge from anonymous market activities. Moreover, they argued that the advantages of established (formal and informal) networks and proximity externalities can geographically overlap in agglomerations. The subsequent table 2.6 summarizes the mechanisms.

Table 2.6. Mechanisms of knowledge acquisition

	static (knowledge transfer)	**dynamic (collective learning)**
formal/ traded relation	anonymous market relations; contract research; consulting; licenses; buying of intermediate goods and knowledge from knowledge suppliers	co-operation/formal networks; link transaction; R&D co-operations with agents at a proximate or distant location; shared use of R&D facilities making repeated and similar transactions with identifiable and distinct partners
informal/ untraded relation	externalities/spillovers; recruitment of specialists; monitoring of competitors; participation in fairs, conferences; reading of scientific literature, patent specifications	milieu/informal networks; informal contacts between agents; social networks in agglomerations; proximity sensitive

Source: illustration from Tödtling *et al.* (2010, 6); see also Johansson and Quigley (2003), Storper and Venables (2004), Johansson (2005) and Capello (2009) for similar concepts.

First and foremost, spillovers of knowledge may originate from unplanned interactions, especially in the context of highly populated areas and city agglomerations (see the previous section). This argument has shifted the attention of researchers towards models of urban growth, which are based on agglomeration economies (Glaeser *et al.*, 1992; Fujita and Thisse, 2003; Malecki, 2010). In this respect, knowledge spillovers constitute an important factor of these agglomeration economies and thus influence co-agglomeration and research clustering. Henderson (1999), e.g., concluded that most external economies stem from information externalities (identical to knowledge spillovers).[194] Accordingly, firms, research institutes and even single researchers within the system produce economically useful knowledge, which can unintentionally spill over to other agents without (full) financial compensation (see section 2.1.6.1).

However, firms that are involved in market transactions have the possibility to take additional advantage of unintended knowledge spillovers. Moreover, the co-operation of firms in sharing the costs (and risks) of an investment may result in a meaningful exchange of valuable knowledge. Nevertheless, this cannot be considered as a pure knowledge spillover, since it is first and foremost an exchange of information (or knowledge) crafted on a market transaction (and monetary flow) (Breschi *et al.*, 2005; Johansson, 2005). Consequently, pecuniary (market-led) flows of knowledge have to be separated theoretically from non-compensated flows (i.e., spillovers that are pure technological externalities) and less-compensated knowledge spillovers (rent spillovers) (Verspagen, 1997).[195] However, such a differentiation is difficult regarding empirical studies.

As no consensus or established approach exists to differentiate between pure spillovers and flows, the empirical analysis in this study is restricted to the identification of research collaboration linkages between agents and regions in a pan-European context, i.e., co-patenting linkages (see chapter 4). It is argued that such linkages represents the only practical measure of knowledge transfer when analyzing hundreds of regions.[196] The extent to which collaboration linkages, e.g., co-patenting linkages, induce spillovers is not challenged in this thesis.

2.1.7.2. Tacit versus Codified Knowledge and the Embodiment Concept

Literature distinguishes between three essential properties of knowledge as an economic good (Nonaka and Takeuchi, 1995; Lissoni, 2001; Foray, 2004): (i) parts of knowledge are non-excludable, which makes it difficult to control or to prevent others from using it; (ii) parts of knowledge are non-rival, which means that other agents can use it, even simultaneously, and therefore it is inexhaustible. However, there is also a discussion related to the

[194] For similar arguments refer to Fujita and Thisse (1996), Black and Henderson (1999b), Black and Henderson (1999a) and Glaeser (2008).

[195] Griliches (1992c) distinguished between "rent spillovers" and true "technological spillovers" (see also Verspagen, 1997). Market spillovers are a synonym for rent spillovers. For an understanding of the conceptualization of rent spillovers, the thesis follows the arguments of Griliches and Jaffe in assuming that some knowledge enters the production function of firms with less compensation, which means that some knowledge is internalized by other firms for a lower price (quality price model). Besides this conceptualization, knowledge spillovers can also enter the production (or utility) without any compensation, which then represents technological externalities.

[196] See Ejermo and Karlsson (2006) for a similar argumentation.

appropriability of knowledge (Cohen and Levinthal, 1990; Capello, 2007; Malecki, 2010);[197] (iii) knowledge is cumulative in nature; although old knowledge becomes partly obsolete as best practice technologies and processes advance, some parts of the knowledge base can remain essential. Accordingly, the three mentioned characteristics seem to be essential when discussing economies of agglomeration and the co-location of agents with respect to knowledge generation and transfer, especially in knowledge-intensive industries.

Nonaka and Takeuchi (1995) have introduced a conceptual framework based upon tacitness and codifiability of knowledge, which helps to classify the transformation of knowledge and implicitly its diffusion into different categories:[198] (i) combination, which is the transfer/transformation of explicit (codified) knowledge to explicit knowledge; (ii) internalization, which is the transfer/transformation of explicit knowledge to implicit (tacit) knowledge; (iii) externalization, which represents the transfer/transformation of implicit knowledge to explicit knowledge via codification; (iv) socialization, that encompasses the transfer/transformation of implicit knowledge to implicit knowledge (Nonaka and Takeuchi, 1995).[199] Channel (iv) is especially of pivotal interest in recent debates on inventor mobility, e.g., the mobility of engineers (Almeida and Kogut, 1999; Breschi and Lissoni, 2009) as will be discussed in the empirical review (see sections 2.2.5 and 2.2.6). The empirical contribution on co-patenting linkages within European inventor networks is, however, mainly focusing on channel (i), although co-inventorship activity is considered to rely on significant parts of implicit knowledge. Moreover, absorption by other agents needs specific technological knowledge and skills.

The approach just described has similarly been proposed by Polanyi (1966). He introduced the differentiation between tacit (implicit) and codified (explicit) knowledge. Stickiness of knowledge is mainly based upon three assumptions: (i) difficulties in exchanging knowledge over long distances, (ii) a context-specific nature that needs common social, organizational and even institutional set-ups, and (iii) the necessity of organized learning processes. Stickiness and thus a substantial increase in the need for geographical, technological, and organizational proximity for economic interaction, and an increasing need for face-to-face contacts (i.e., "handshakes"), may in particular be useful to explain persistent research clustering and agglomeration of innovative activity (von Hippel, 1994; Audretsch, 1998; Feldman, 2000).[200] In the same line of reasoning Gertler (2003, 79) argued that

> "[w]hen one combines these two features of the innovation process - the centrality of sticky, context-laden tacit knowledge and the growing importance of social interaction - it becomes apparent why geography now matters so much."

---

[197] Appropriability and the concept of absorptive capacity are related to knowledge accumulation. Cohen and Levinthal (1990, 129) argued that *"research on memory development suggests that accumulated prior knowledge increases both the ability to put new knowledge into memory, what we would refer to as the acquisition of knowledge, and the ability to recall and use it."*

[198] According to Nonaka and Takeuchi (1995, 59), *"[t]acit knowledge is personal, context-specific, and therefore hard to formalize and communicate."*

[199] For further discussions of codification processes from implicit to explicit knowledge refer to Cowan *et al.* (2000), Feldman (2000), Fischer (2001), Scherngell (2007), Jensen *et al.* (2007), Cooke (2007).

[200] von Hippel (1994, 432) has argued that *"[w]hen information transfer costs are a significant component of the cost of the planned problem-solving work, it is reasonable that there will be a tendency to carry out innovation-related problem-solving activity at the locus of sticky information."*. See also Breschi *et al.* (2005) and Cooke (2007).

Tacitness of knowledge is context-specific. Some people may find it simple to articulate such (pieces of) knowledge, others, however, do not (Cowan *et al.*, 2000; Lissoni, 2001; Johansson and Quigley, 2003). In this respect, epistemic communities (of science) have developed their own language, institutional set-up and communication codes for codifying, transmitting and securing scientific knowledge and for reinforcing tacitness and reducing externalities (Steinmueller, 2000; Lissoni, 2001; Balconi *et al.*, 2004). Accordingly, knowledge can be considered a freely available good within each epistemic community (club good). Related to the degree of codifiability of knowledge, the subsequent table 2.7 summarizes the different modes of knowledge diffusion.[201] The table particularly centers the type of knowledge (tacit versus codified) and possible working channels of knowledge transmission.

Table 2.7. Modes of transfer of tacit and codified knowledge

Acquisition mechanism	**Codified knowledge/ public good**	**Codified knowledge/ private good**	**Tacit knowledge**
non-market acquisition/ informal networks	learning and absorption of specific knowledge by studying documents, data, blueprints; education and graduation from organizations	externalities from reverse engineering; strategic brain drain; studying patent descriptions	externalities from job hopping; learning by doing, watching, interacting; inventor networks
market acquisition and/or formal networks	-	purchase of technology (anonymous market); licensing of technologies protected by patents; software R&D assignments; M&A	acquisition of researchers and engineers as carriers of tacit knowledge; established, stable and repeated co-operations and R&D networks as carriers of tacit knowledge

Source: illustration taken from Franz (2010, 8); see also McCann *et al.* (2002), Johansson and Quigley (2003), Johansson (2005), Scherngell (2007) and Jensen *et al.* (2007) for similar concepts.

To conclude, possible transfer channels (and carriers) of knowledge spillovers are summarized in table 2.8 for a complementary review. According to the table, the working channels can be classified with respect to the "embodiment" issue. The table distinguishes between knowledge embodied in people and knowledge enclosed in goods. It is worth noting that the empirical analysis draws attention to the distribution of research activity, where the focus is on the distribution of patenting activity, inventors and inter-regional co-patenting networks, but not explicitly on people, their social networks or their migratory movements.

201 For further discussions and definitions of tacit knowledge refer to Feldman (1994b), Audretsch (1998), Feldman (1999), Ottaviano and Thisse (2000), Lissoni (2001), Asheim and Gertler (2005), Döring and Schnellenbach (2006), Scherngell (2007), Jensen *et al.* (2007) and Franz (2010).

Table 2.8. Transfer channels of knowledge via agents, goods and documents

Embodied in agents	Embodied in products/documents
mobility of labor; especially mobility of highly skilled people via job hopping	technology white book and scientific publications
mobility of labor via entrepreneurship and spin-offs	patent documents, patent application, patent licensing
mobility of labor via conferences, expositions	vertical linkages/technology transfer via reverse engineering of intermediates (rent spillovers)
(formal) inventor networks and informal social networks	horizontal linkages/technology transfer via reverse engineering of final goods (rent spillovers)

Source: illustration based on Feldman (2000), Lissoni (2001), Johansson and Quigley (2003), Johansson (2005), Scherngell (2007) and Capello (2009).

2.1.7.3. Agglomerations, Innovative Milieus and the Proximity Hypothesis

Networks are considered to differ from agglomerations although they may have some overlaps (Burger *et al.*, 2009). The formation and efficiency of an agglomeration arises from its quasi public good character (non-rivalry, non-excludability). Agents, households and firms within an agglomeration share the benefits of spatial proximity; only spatial distance makes it a "quasi club good" if agglomeration economies show strong distance decay effects. In contrast, an economic network between agents represents some kind of "private capital" which originates from individual investments that are shared solely by network participants. Economic networks emerge from collective decisions which are made by groups which run private institution. However, agglomerations and industry clusters rely mostly on public institutions (Johansson and Quigley, 2003; Johansson, 2005).

According to Johansson (2005), networks are generally established in order to facilitate the exchange of assets within and between organizations, regions and countries to reduce transaction costs. The transaction cost approach (TCA) is helpful to determine whether markets, organizations or a combination is more efficient in coordinating exchange of assets. Intangible economic networks represent transaction agreements and routinized arrangements between agents and firms; they differ from the anonymous market and physical networks (see also Williamson, 1975, 1979; Johansson and Quigley, 2003; Wilhelmsson, 2009). An economic network can be regarded as an organization of interlinked agents which combines features of the market and the firm. The network internalizes some interaction costs and is built on several agreements, similar to standard market contracts. Accordingly, networks exist in order to reduce transaction costs. Regarding research and co-inventor networks, it is generally argued that long-distance research collaborations are costly. When transactions are generally distance-sensitive, persistent transaction linkages can overcome spatial distance and reduce costs. However, there are always fixed costs associated with the process of establishing a network between agents and/or firms. The transaction costs may be lower inside an agglomeration, but long-distance linkages also reduce costs if they are persistently used (McCann *et al.*, 2002; Johansson and Quigley, 2003; Johansson, 2005). Accordingly, network linkages can emerge within agglomerations in different forms, e.g., industrial complexes and social networks, but also between regions and agglomerations (McCann *et al.*, 2002; Johansson and Quigley, 2003; Burger *et al.*, 2009). The co-patenting analysis in chapter 4 builds upon this idea, as R&D collaboration is regarded as a meaningful channel for knowledge flows (Ejermo and Karlsson, 2006; Burger *et al.*, 2009).

Contributions to regional knowledge networks and the innovative milieu since the late 1970s combined an "industrial," "spatial" and "cognitive" dimensions of agglomeration economies (Capello and Faggian, 2005; Capello, 2009). An interesting debate in the context of spatial proximity was related to the distance decay effects of knowledge transfer due to the aforementioned properties of knowledge. Although geography was included in the "spatial" and "industrial" concepts, the micro-foundations of knowledge transmission remained a "black box" (Lissoni, 2001; Breschi and Lissoni, 2001a; Storper and Venables, 2004). However, a salient feature reported by many empirical studies is that greater distance tends to decrease the frequency of economic activities and interactions among observations, especially with respect to knowledge transmission and research activity. For this reason, intellectual tasks are considered to be influenced enormously by geographic distance (Maggioni *et al.*, 2007; Maggioni and Uberti, 2009; Hoekman *et al.*, 2009). That being the case, the observed patterns of spatial knowledge diffusion are said to originate from the different properties and attributes of knowledge as described above. In light of this, scholars regularly make use of the tacit knowledge concept in order to explain and categorize the relationship between spatial proximity and knowledge transmission. Thus, the tacit knowledge concept is linked with the social network approach (Lissoni, 2001; Breschi and Lissoni, 2003; Storper and Venables, 2004).[202] Such (social) networks are generally related to the work of, e.g., Granovetter (1973) and can be regarded as a response to the hierarchies model of Williamson (1975) (see also Johansson, 2005; Capello, 2009). The social network approach argues that mutual trust relations between agents in different organizations are at least as important as decision-making hierarchies within the individual organizations (McCann *et al.*, 2002).

Opposed to the commonly applied a-spatial public good character of knowledge in early endogenous growth models (Romer, 1986; Jones, 2004), the French "proximity school" considers localization as a relevant factor due to the costly transmission of knowledge across space as it is embedded, non-codified and not explicitly stated. For this reason, it is not easily transferable and its exchange is extremely costly and sensitive to the social context, which is itself a local phenomenon (localized social networks). Such localized interdependencies create a milieu effect (Camagni, 1991b,a; Capello, 2009). The approach explicitly supports the observed phenomena of spatial concentration and clustering of research activity (Paci and Usai, 2000b; Balconi *et al.*, 2004; Maggioni *et al.*, 2007). The local "stickiness" of parts of knowledge is assumed to support localized intra-regional knowledge-intensive interaction (Nonaka and Takeuchi, 1995; Lissoni, 2001; Johansson and Quigley, 2003).[203]

[202] Similarly, Lundvall (2007, 103) suggested that *"[a] key difference between firms, sectors, regional and national systems is the role played by respectively codified and tacit knowledge in the innovation process."*

[203] Refer also to Scherngell (2007), Lundvall (2007) and Powell and Giannella (2010).

Regarding spatial distance between epistemic communities, it is argued that it is difficult to communicate and collaborate over considerable distances, as distance generally increases costs of establishing and maintaining knowledge access (Balconi *et al.*, 2004; Hoekman *et al.*, 2009). Long-distance transaction and its coordination are generally more time-consuming, cost-intensive and hinder a productive dialogue compared to collaborations at a proximate distance. According to that, the efficiency of knowledge exchange at a distance is a function of codifiability, teachability and complexity (Breschi and Lissoni, 2003; Storper and Venables, 2004; Capello, 2009). Related to these interdependencies, Storper and Venables (2004), Bathelt *et al.* (2004) and Maskell *et al.* (2005), among others, argued that local systems are determined by several dimensions of proximity. They emphasized that localized learning processes are heavily influenced by learning by interacting (vertical), learning by monitoring (horizontal) and some neighborhood-effects and "local buzz" (social dimension), which makes the city a central place. Similarly, Glaeser *et al.* (1992, 1126) argued that

> "[i]f geographical proximity facilitates transmission of ideas, then we should expect knowledge spillovers to be particularly important in cities. After all, intellectual breakthroughs must cross hallways and streets more easily than oceans and continents."

Obviously, the above presented concept of "knowledge tacitness" is used and extended by different schools of thought; e.g., economic geographers (Boschma and Frenken, 2009; Hoekman et al., 2010), economists (Fujita and Thisse, 1996; Ottaviano and Thisse, 2001; Baldwin and Martin, 2004; Gosens and de Vaal, 2010) and innovation scholars (Audretsch, 1998; Audretsch and Feldman, 1999; Feldman, 1999, 2000; Breschi and Lissoni, 2003).

The presented assumptions and concepts lead to the conclusion that the region represents a central place of knowledge production, recombination and sharing (and of knowledge diffusion in general).[204] This idea is considered to be especially relevant with regard to research activity that is in general spatially concentrated in large cities and urban regions (Fischer, 2001; Cooke, 2001; Henderson, 2010).[205]

Another working channel of knowledge exchange at a proximate distance are spin-offs, which are considered to be located in the neighborhood of their parent organizations (or competitors), which increases spatial concentration and thus research clustering (Zucker *et al.*, 1998; Audretsch *et al.*, 2005; Ponds *et al.*, 2010). According to Audretsch and Feldman (2004, 2733), spin-offs locate in proximity to existing clusters as

> "[s]uch start-ups typically do not have direct access to large R&D laboratory. Rather, these small firms succeed in exploiting the knowledge and experience accrued from the R&D laboratories with their previous employers."

To conclude, spatial proximity affects network formation and place makes a difference (Glückler, 2007). Spatial proximity between agents and the emergence of milieus facili-

[204] Storper (1997, 44) concluded that the region itself represents *"[a] site of important stocks of relational assets."*

[205] This has also been issued by Lagendijk (2001, 81) who argued that *"[d]ue to the changes in industrial organisation, notably the rising importance of networking and learning, the region has come to be seen as a highly appropriate level for knowledge production."*

tate both unintended knowledge spillover and intended knowledge exchange (in localized networks). Moreover, co-location is considered to enable "collective learning" within the boundaries of the territory (Capello and Faggian, 2005; Johansson, 2005; Capello, 2007). In this respect, network economies and agglomeration economies may overlap in urban areas and innovation clusters as soon as informal relationships (social networks) are present. If the agents are located in a single location and the effects of local interdependencies are low, then there is no clear-cut distinction between the relationships based upon network linkages and the formal contracts of an anonymous market transaction. In general, it is suggested that spillovers in agglomerations and spillovers in networks lead to a more rapid development in technology, innovative capacity, innovation output and productivity (Johansson and Quigley, 2003; Capello, 2007, 196). However, network linkages differ from anonymous market interactions in case of long-distance interactions as will be discussed in the following section.

2.1.7.4. Long-Distance Linkages and R&D Collaboration Networks

Another relevant type of economic and social relations represents trans-territorial cooperation agreements, i.e., inter-regional network linkages. Continuous and stable networks originate from the complexity of strategic alliances between agents (von Hippel, 2005; Ejermo and Karlsson, 2006; Powell and Giannella, 2010). It is generally assumed that transaction networks facilitate knowledge exchange between agents. Formal inter-regional network linkages are considered to function as a substitute for spatial proximity in the R&D process (proximity externalities in agglomerations) and in the knowledge-transmission process, respectively (Foray and Steinmueller, 2003; Johansson and Quigley, 2003; Porter *et al.*, 2005). According to Johansson (2005), advantages emerge from both transaction and innovation externalities if the network linkages show continuity (see section 2.1.6.5) what Capello (2009) calls the "synergy dimension" of agglomeration economies.

Related to the analysis of research networks and patenting activity, Maggioni *et al.* (2007, 475) recently argued that

> "[i]t is almost impossible to disentangle the tacit from the codified elements of knowledge flows running across European regions. It is reasonable however, to assume that, since tacit knowledge needs face-to-face contacts and these contacts are inversely related to geographic distance, long-distance relations imply a greater role played by codified knowledge than the relationships between nearer regions."

Therefore, a pivotal role is attributed to inter-regional research linkages and the emergence of trans-territorial networks as they are said to complement the locational advantage in agglomerations and milieus. This study follows this argument in analyzing European co-inventorship linkages in a spatial and technological perspective. For clarification purpose, table 2.9 offers a taxonomy of transaction-link based externalities. These externalities represent idiosyncratic inter-firm relations, which provide benefits that emerge from outside the ordinary (anonymous) market, i.e., a quasi-market setting (Johansson, 2005, 120). It is argued that a research network shares the extra-market properties of a club (Capello, 2009). Opposed to market anonymity, link transactions represent repeated and similar transactions between identifiable (non-anonymous) and distinct partners and build upon

Table 2.9. Network linkages and externalities

Externality type	Type of transaction	Principle mechanism
Network	Supply-chain externalities and complex network (transaction) externalities	Agents belonging to a network form a kind of club, through which they have accessibility to joint assets that facilitate transactions.
	Knowledge externalities within networks via transaction linkages	Knowledge spillovers are possible between agents that belong to the same network.
Vertical	Upstream or input- cost link externality. Established, repeated transactions.	A persistent (network) link between an input supplier and a customer reduces transaction costs; linkages may be distance sensitive (setup costs); positive transaction externality for the customer.
	Downstream or delivery-cost link externality. Established, repeated transactions.	A persistent (network) link between an input supplier and a customer reduces transaction costs; linkages may be distance sensitive (setup costs); positive transaction externality for the supplier.

Source: illustration based on Johansson (2005, 121); see also Johansson and Quigley (2003) and Capello (2007, 2009).

prior interactions. Moreover, investments in network linkages (i.e., sunk costs) affect future interactions and are thus profitable if linkages are expected to be used permanently (Johansson and Quigley, 2003; Johansson, 2005; Bahlmann *et al.*, 2009).[206] The partners in inter-regional networks are always selected single economic units (researchers, firms) and geography represents only one dimension that affects the identification and selection of potential partners (Capello, 2007, 198-199).

To conclude, an established and stable network can reduce the "effective" distance between agents located in different regions in reducing transaction costs. In the case of infeasible co-location, networks have potentialities to act as a substitute for co-location and co-agglomeration and offer advantages similar to innovative milieus, especially if the progress in communication technologies is fast and the organizational structures improve in the course of time. Transaction linkages via networks have positive effects (Johansson, 2005). First, they are defined by reducing coordination and search costs and the set-up of interactions with identical agents induces sunk costs. Second, they are non-anonymous and repeatable, which reduces uncertainty. Third, knowledge externalities can occur as a by-product of network transactions in a proximate distance or long-distance relationship (innovation externalities).

206 Refer to the transaction cost approach of Williamson (1979).

2.1.7.5. Localized Networks versus Inter-Regional Network Linkages

As has been argued in the previous sections, the relationship between agglomeration economies, local interaction, innovative capacity and economic growth represents an established and growing field of research. With respect to proximity and networks (sections 2.1.7.3 and 2.1.7.4) it is suggested that agglomerations have potentialities to develop new knowledge and technologies faster and to show higher productivity growth (Ellison and Glaeser, 1997; Johansson and Quigley, 2003; TerWal and Boschma, 2009).[207] Information and knowledge diffuse quite quickly among agents who belong to the same social and/or formal network, i.e., the same transaction network and/or epistemic community, especially when they co-locate at a proximate distance (proximity externality). However, it has also been argued by economists and geographers that well-functioning (formal and informal) networks between agents may be able to substitute for spatial proximity in the context of research activities and the processes of intended (and unintended) knowledge diffusion (Johansson and Quigley, 2003; Bergman, 2009; Burger et al., 2009).

However, the presented frameworks and remarks on agglomeration economies in the former sections solely focus on intra-regional factors. Spatial interaction is implemented as a "black box" of spillovers, if at all. Regarding this issue, Harris (2008, 22) complained that

> "[r]esearch on agglomerations/clusters has focused on the internal characteristics and mechanisms in those places and diverted attention from the necessary distinct, even global, linkages that competitive places require."

Spatial distance is commonly considered a general barrier to interaction and knowledge diffusion, meaning that many frameworks have not taken into account inter-regional linkages and research collaborations via network pipelines. Increasing attention in recent theoretical (and empirical) contributions to economic geography (TerWal and Boschma, 2009; Neffke *et al.*, 2009; Boschma and Frenken, 2010) and geographical economics (Fujita and Thisse, 1996, 2003; Gosens and de Vaal, 2010) is drawn to the question of whether geography is still the dominant factor for productivity and competitiveness and for regional development and economic growth (see section 2.1.7.3). Inter-regional linkages may represent a substantial factor for regional innovative capacity and growth (see section 2.1.7.4).

Geographers differentiate between the "space of places" and "space of flows" (Castells, 1996; Burger *et al.*, 2009).[208] The former concept supports the idea that location and intra-regional factors matter for knowledge transfer and innovative capacities, whereas the latter concept centers on the assumption that networks are the pivotal factor. In light of this discussion, agglomeration economies, inter- and intra-regional networks and long-distance research collaborations are considered pivotal factors that likewise operate and affect the geography of innovation. This theoretical consideration has meaningful effects on research methodologies and agendas with respect to (i) firm-level network analysis, (ii) intra-cluster and intra-regional studies, (iii) inter-cluster and inter-regional studies and

[207] In this respect, the debates also cover the questions whether sectoral specialization or a more diversified regional structure is fruitful for knowledge generation, diffusion and growth in employment, productivity, and research output in general (section 2.1.6). Moreover, researchers are interested in the question if distance decay effects of knowledge diffusion may have essential influence (sections 2.1.7.1 and 2.1.7.2).

[208] See also Boschma and Lambooy (2002) and TerWal and Boschma (2009).

(iv) research on agglomeration economies (TerWal and Boschma, 2009; Burger *et al.*, 2009; Maggioni and Uberti, 2009). Besides being a crucial theoretical debate, it is moreover an empirical discussion to which this thesis will contribute (see chapters 3 and 4).

From a theoretical perspective, Breschi and Lissoni (2001a, 2003) criticized the existing foundations of knowledge transmission by arguing that it is not geographical proximity itself (and distance decay) that generates localized knowledge diffusion (i.e., spillovers) and thus research clustering. They argued that the central driving force is networks of agents and firms, i.e., the mobility of entrepreneurs and inventors across firms and regions in specific technology fields, which may tend to be geographically localized (section 2.1.7). This consequently might cause knowledge externalities to exhibit strong distance decay effects and knowledge flows and collaborations to have a limited spatial distance because the underlying knowledge infrastructure, i.e., inventor and research network, is assumed to be confined to the boundaries of the cluster (or region) (Breschi and Lissoni, 2001b; TerWal and Boschma, 2009; Breschi and Lissoni, 2009).[209] According to this point of view, clusters, innovative milieus and industrial districts are considered to enclose mostly all relevant parts of the research network (see section 2.1.7.3). However, such an inward perspective is likely to ignore a meaningful fraction of linkages which reaches beyond regional borders (Boschma and Lambooy, 2002; TerWal and Boschma, 2009; Burger *et al.*, 2009).

2.1.7.6. City Networks and Inter-Regional Research Collaborations

Fading out the local boundaries of innovative milieus (and regions, respectively), there is significant evidence in the literature that inter-regional linkages may represent increasingly important factors with respect to knowledge generation, transfer and diffusion (Bathelt *et al.*, 2004; Capello, 2007; TerWal and Boschma, 2009) - at least since the last wave of globalization has induced additional fragmentation and offshoring of R&D activities (Dicken, 2000; Belitz *et al.*, 2006; Legler and Krawczyk, 2006).[210] Thus, knowledge is assumed to be intentionally transmitted in research networks, which are structured via contractual agreements (see section 2.1.7.4). Accordingly, the ICT revolution is said to have created knowledge infrastructures, such as pipelines that allow for trans-border knowledge flows, and which facilitate the search, combination and recombination of different types of knowledge and information beyond the boundaries of a location (Steinmueller, 2000; Johansson and Quigley, 2003; Bathelt *et al.*, 2004).[211]

In economic geography, several researchers have discussed a globalization driven "death of distance" and the transformation of the globe into a "flat world" (Cairncross, 2001; Dicken, 2000; Maggioni and Uberti, 2009).[212] On the other extreme, the traditional agglomeration economies literature, as presented and discussed in the previous sections, clearly leaves out the possibility of significant knowledge flows between research clusters via persistent

[209] For an overview refer to Capello (2007), Burger *et al.* (2009) and Bergman (2009).
[210] For a detailed overview refer to Capello (2007).
[211] Additionally, some authors see the benefit from global knowledge pipelines in overcoming trajectories and potential lock-in (Asheim and Isaksen, 2002; Kilkenny, 2010).
[212] In 1995, the Economist raised meaningful doubts that geographic location is still a central factor for the innovation process in a world full of email, fax machines and the internet ("The Death of Distance," The Economist, 30 September, 1995). See also Audretsch (1998).

inter-regional co-inventor networks. In view of this, many researchers argue that distance and clustering are persistent phenomena, even in the internet era (Gertler, 2003; Scott and Storper, 2003; Crafts and Venables, 2003).

An interesting concept, related to the just-described network linkages and research collaborations, is the conceptual paradigm of city networks that can be considered an alternative interpretation of existing transaction linkages within and between metropolitan regions (Capello, 2007, 78). Cities could be regarded as spatial entities in an urban system of horizontal and vertical relationships. Urban systems are formed by networks which consist of specialized but complementary (production) centers and specific input-output relations which lead to inter-regional transactions. Economies of scale depend on sectoral specialization and fragmentation of tasks. Horizontal transaction linkages between different places create synergy networks between cities. Such transaction networks enable economies of scale as markets are interconnected, i.e., network externalities emerge for network members (Capello, 2007, 79). Furthermore, innovation networks between cities, i.e., research networks (see section 2.1.7.4), are considered to connect regions. Such inter-regional co-inventor networks emerge from persistent research linkages at the firm level (Balconi *et al.*, 2004; Maggioni and Uberti, 2009). Economic relations between metropolises/cities based upon cooperative (research) linkages may enable some kind of economies of scale that do not depend on the size of the regional economy. However, the efficiency of the linkage depends on the economic and technological features of the network (and set-up costs) (Johansson, 2005). Persistent inter-regional linkages could generate economies (or externalities) of specialization (via division of labor), co-operation, synergy and/or innovation (Capello, 2007; Maggioni and Uberti, 2009). In the context of R&D activities, urban systems connect knowledge hot spots via network linkages that represent (persistent) knowledge pipelines (Bathelt *et al.*, 2004; Powell and Grodal, 2005; Bahlmann *et al.*, 2009).

Belitz *et al.* (2006) discussed the emergence of inter-regional research collaborations under the well-known label "the internationalization of industrial R&D." They argued that companies generally intensify their foreign research activities as they need access to the forefront knowledge that is located in global centers of excellence, i.e., research clusters. Moreover, Belitz *et al.* (2006, 59) concluded that

> "the more knowledge-intensive technology fields are, the closer are research activities to the scientific forefront of knowledge."

Research networks are considered to be increasingly important as they contribute to research output of superior quality (Stephan, 1996; Wilhelmsson, 2009). Accordingly, tendencies of inter-regional and international network formation and research clustering are said to co-exist.

In light of the debate on R&D internationalization, Malecki (2010) similarly argued that R&D is no longer confined to firms' R&D laboratories (see also Powell and Giannella, 2010). Today, research takes place within dispersed networks, both internal and external to the single firm and region (Bergman, 2008; Burger et al., 2009). Moreover, Malecki stated that global networks of production and innovation contain widespread connections among sources of knowledge (Ernst, 2002; Ernst and Kim, 2002), which is in line with the "new ecology of R&D movement," "open innovation models" and "collective invention approaches" (Johansson and Quigley, 2003; Chesbrough, 2003; Powell and Giannella, 2010). Collective invention enables the access to information (and knowledge) and the upgrading of technological skills at the same time. Accordingly, the (inward) R&D-based regional knowledge production function seems to be vanishing and represents an increasingly incorrect approximation of reality (Coombs and Georghiou, 2002; Malecki, 2010). The internationalization of production induces at the same time an internationalization of R&D tasks. It is argued that the fragmentation and specialization of R&D activities has let to a more international generation and exploitation of technological and scientific knowledge in a co-evolutionary process (Belitz *et al.*, 2006; Frietsch and Schmoch, 2006; Malecki, 2010).

Bathelt *et al.* (2004) argued that knowledge, which enters the cluster via global research pipelines, is likely to spill over to other agents. Accordingly, the actors' regional research network is the key in understanding knowledge diffusion. As a consequence, the "global knowledge pipeline story" is interpreted as an essential progress in the common cluster literature, as it turns the view of clusters as bounded regions into a (dynamic) network perspective (Bathelt *et al.*, 2004; Bahlmann *et al.*, 2009; Burger *et al.*, 2009).[213] Similarly, Bresnahan *et al.* (2001) pointed out the necessity of inter-regional linkages for the rise and development of clusters (see also Breschi and Malerba, 2001).

TerWal and Boschma (2009) argued that the inflow of a large variety of knowledge through inter-regional research linkages compensate historical lock-ins and relative specialization into mature industries. Similarly, Boschma and Iammarino (2007) suggested that inter-regional research collaborations and inflows of knowledge might particularly enhance and upgrade regional knowledge bases (see also Asheim and Isaksen, 2002; Bathelt *et al.*, 2004; Bahlmann *et al.*, 2009).

Different benefits of research collaborations have been discussed in the literature: (i) better access of firms to information and knowledge (i.e., knowledge bases); (ii) intense linkages and co-operation between agents leading to the accumulation of skills; (iii) higher response capacity, stimulating effects and sources of creativity; (iv) reduction of risk and moral hazard; (v) reduction of search and transaction costs; (vi) trust-based relationships and formal linkages, social cohesion; and (vii) enhancing the potential visibility of the single agent (Fraunhofer, 2009).[214]

Mattsson *et al.* (2008) grouped the above-mentioned advantages and benefits into four categories: (i) financial reasons (e.g., access to funding and sharing facilities); (ii) social factors (e.g., networking and acknowledgements within communities); (iii) a preference for working in teams and not in isolation; and (iv) improving technical, analytical and theoretical knowledge and political factors (including framework programs to ease collaboration). Powell and Giannella (2010) discussed these benefits under the label "collective invention."

An increase in inter-regional research collaboration, i.e., co-inventor linkages, could be related to an advancing regional integration that is based upon falling transportation costs, increased mobility of researchers and employees, footloose entrepreneurs, communication technologies that facilitate codified knowledge transfer and the transformation of tacit

[213] Johansson and Quigley (2003) label this development the "agglomeration economies vs. networks debate."

[214] See also Hotz-Hart (2000), Bergman (2009) and Burger *et al.* (2009).

knowledge to more codified forms. Accordingly, the spatial distance of network relations between agents is considered to have increased over time as the knowledge base of an industry becomes increasingly codified (Lissoni, 2001; Gertler, 2003; TerWal and Boschma, 2009).[215]

It is also argued that the different stages in a cluster and/or industry life cycle may be interrelated with the frequency and intensity of inter- and intra-regional knowledge flows via co-inventor networks. In this respect, it is suggested that inter-regional co-inventor linkages are essential for both cases, for regions dominated by mature industries and/or emerging regions that have to build upon external knowledge. Agents in regions benefit from inter-regional flows of knowledge and regions experience a significant upgrade of their specific knowledge base (Bathelt *et al.*, 2004; Singh, 2005; Powell and Giannella, 2010). Nevertheless, it is argued that there is a meaningful difference between a "weightless" and "spaceless" economy.[216]

2.1.7.7. Agglomeration vs. Networks: Critical Remarks

With respect to the above-listed carriers and working channels of knowledge, Audretsch and Feldman (2004) and Capello (2007, 2009), among others, discussed the lack of a clear microeconomic conceptualization (foundation) of transfer and diffusion channels in the majority of knowledge spillover studies. Similarly, Breschi and Lissoni (2001a) reviewed recent studies and offered a critical discussion of the knowledge spillover approach in general and the applied econometric approaches. Although knowledge spillovers are said to perform a pivotal role in the seminal theoretical frameworks, especially in the 1990s' and 2000s' workhorse models of endogenous growth (section 2.1.6.6 and 2.1.6.7), there is still a serious gap in the literature from a theoretical (and empirical) point of view. This is especially a severe issue in knowledge production function approaches, which introduce spatial lags of R&D activity or innovation output. However, such approaches do not give any indication about the knowledge diffusion channels, which turns them into a "black box" (Breschi and Lissoni, 2001a; Döring and Schnellenbach, 2006).[217]

Boschma and Frenken (2006) distinguished between different working channels of knowledge spillovers: (i) imitation, (ii) spinoff firms, (iii) social networks, (iv) labor mobility and (v) R&D collaborations via collaborative networking (see also Burger *et al.*, 2009). In this respect, several authors consider networks as meaningful sources of knowledge spillovers and flows (Bergman, 2009; TerWal and Boschma, 2009). However, regarding knowledge-intensive industries, it seems rather impossible and counterproductive to separate the agglomeration economies debate from network approaches; in particular when agglomeration economies are assumed to stem from knowledge transmission via intra- (and inter-) regional

[215] A high degree of proximity may increase the probability of agents getting connected to others; however, it is expected that the effects of network linkages on innovative activity are rather ambiguous (TerWal and Boschma, 2009; Boschma and Frenken, 2009a).

[216] Geppert and Stephan (2008, 209) argued that *"[t]he internet and knowledge society may increasingly become weightless, but there is no indication that it is also becoming spaceless."*

[217] For further discussions refer to Feldman (2000), Audretsch and Feldman (2004), Capello and Faggian (2005) and Feldman and Kogler (2010).

research networks (see chapter 2, section 2.1.7.5). From a conceptual point of view, knowledge sourcing via the market (buying from knowledge suppliers) and knowledge transfer in inter-regional research co-operations via research networks (transaction link externalities) represent an alternative to the locational advantage in agglomerations.

A region's innovative capacities are affected from two sides. On the one hand, collective learning processes and knowledge transmission, which are mainly based upon local interaction and the transmission of tacit knowledge in local networks, are seen as pivotal factors of co-location and research clustering. Researchers assume that knowledge externalities indeed happen due to spatial proximity (proximity externalities), institutional thickness of the regional system and a significant tacitness of knowledge (Johansson, 2005; Cooke, 2007; Tödtling *et al.*, 2010). On the other hand, inter-regional flows of knowledge originate from anonymous market transactions and from continuous inter-regional research networks (network linkages) that connect global knowledge hot spots (Capello, 2007). Moreover, the existence of trans-territorial networks supports the idea of a technology-field-specific minimum degree of openness and "absorptive capacity" of regional systems and clusters.[218]

First and foremost, elite researchers (and their research locations) take advantage from inter-regional research linkages, who work at the cutting edge (in epistemic communities) (Hoekman et al., 2010). Under the assumption that elite researchers and research centers are much more spatially concentrated, higher shares of inter-regional linkages are expected to exist and emerge within this group of researchers but not within less advanced ones (Capello, 2007; Frenken et al., 2007; Hoekman et al., 2010). Elite researchers, who are mainly located in dense urban areas, may primarily build collaboration linkages with other leading research centers but not with less advanced ones, which then form an urban system of city networks (Capello, 2007, 2009). However, backward research locations may also profit from inter-regional collaboration networks as they become connected to other knowledge hot spots and research clusters (Bathelt *et al.*, 2004; von Hippel, 2005; Maggioni *et al.*, 2007). With respect to the costs associated with such research collaboration linkages, two major improvements that stem from technological progress are observed: (i) decreasing transportation costs (and travel times) and (ii) a significant improvement in information and communication technologies; both developments are assumed to have hampered the effects of spatial distance on inter-regional and international research collaborations (and economic integration in general). In addition, stable network linkages are different from anonymous market transactions as they induce set-up costs.

However, even today, research activities may still face the issue of search and coordination costs in a geographical context (Hoekman *et al.*, 2010). Agents (and research centers) have to bridge geographic distances, which makes research collaboration activities sensitive to physical distance (section 2.1.7.3). As the embeddedness of researchers (and their laboratories and facilities), their mobility, different institutional settings and linguistic areas, differing labor markets and technological specialization (among factors) generally show

[218] Absorptive capacity is a core component of regional systems; however it is assumed to vary between regions and countries (Cohen and Levinthal, 1990). This idea is also formulated by Camagni (1991b, 8), among others, who stated that *"[t]echnological innovation [...] is increasingly a product of social innovation, a process happening both at the intra-regional level in the form of collective learning processes, and through inter-regional linkages facilitating the firms access to different, though localised, innovation capabilities."*

a local nature, most research collaborations are said to happen at a proximate distance (Maggioni *et al.*, 2007; Breschi and Lissoni, 2009; Paci and Usai, 2009). From an institutional economics point of view, it is also argued that regions that belong to different countries are institutionally different and more reluctant to collaborate in research activities (Hoekman *et al.*, 2009). These arguments are completely in contrast to the suggested "death of distance" argument (Castells, 1996; Cairncross, 2001).[219] It can be argued that the intensity of inter-regional collaborations is a function of costs and that collaborations are negatively affected by spatial distance. To optimize costs, researchers have a strong incentive to collaborate with other groups of researchers at an institutionally and spatially proximate distance, i.e., within an epistemic community within a cluster or region. In a European context, one should expect to find highly localized networks.

Moreover, it can be concluded from the above-described concepts, models and debates that agglomerations and the co-location of research activity are pivotal when the knowledge transfer follows unintended patterns (technological externalities) or happens intentionally in spatially concentrated networks. However, knowledge transmission is also said to happen intentionally via inter-regional research linkages irrespective of spatial distance. Although it is impossible to separate clearly tacit from codified knowledge flows, it is, however, reasonable to assume that knowledge transfer is inversely related to spatial distance. Furthermore, long-distance relations imply a greater role played by codified pieces of knowledge.

In summary, the theories described in the previous sections have highlighted central issues related to research clustering, agglomeration economies and inter-regional research networks. The empirical analysis in this study aims to follow explicitly the aforementioned arguments and concerns. In the first part, the empirical analysis in this study will shed light on the regional disparities and spatial concentration of patenting activity in Europe at the regional level (chapter 3). In the second part, the analysis will also emphasize the structures and dynamics of inter-regional research collaboration linkages, i.e., European co-patenting networks at the regional level (chapter 4). On account of this, co-inventor network analysis, among other approaches, can apply EPO patent applications as relational data, as will be discussed in the following part of the survey.[220] The subsequent section 2.2 briefly presents empirical results relating to research clustering, regional R&D spillovers and inter-regional co-patenting networks.

2.2. A Survey of the Empirical Literature

2.2.1. The Co-Evolution of Different Strands of Empirical Research

The empirical literature on the geography of innovation, research clustering and knowledge transmission can be generally classified into six streams: (i) research concentration and clustering studies, which focus on regional disparities and the distribution of research activities; (ii) the knowledge production function approach (KPF), which analyzes the

[219] See also Dicken (2000) and Giddens (2000).
[220] Alternative data sources for network analysis are, e.g., co-publication data (Hoekman *et al.*, 2009; Ponds *et al.*, 2010) and European Framework Programme data (Maggioni *et al.*, 2007).

input-output relationship between R&D and patenting activity in a geographic context; (iii) studies on localization and urbanization economies and regional development; (iv) the patent citation or "paper-trail" approach, which explores the "real" knowledge spillovers and the geographic and/or technological relatedness of forward and/or backward patent citations; (v) studies on the mobility of individuals and their social networks, especially of researchers, engineers and other highly skilled individuals in a spatial context; and finally (vi) studies on inventor networks, research collaborations and co-patenting networks, which use data on intra- and inter-regional research linkages from co-patenting data.

The major contribution of the mentioned research streams is to approach and work out empirical issues related to the distribution of research activity, the causes and effects of research clustering, the functional boundaries of local systems and clusters and the peculiarities of inter- and intra-personal and inter- and intra-regional knowledge transmission. It can be argued that the strands of empirical research combine an "industrial," "geographic," "technological," "socio-cultural" and "cognitive" dimension of research clustering, agglomeration economies and knowledge transmission.

The following is a detailed empirical literature review of the mentioned strands of research in the context of European research clustering, R&D spillovers, the localization of patent citations, local inventor networks and inter-regional network linkages. Although the empirical analyses in this thesis place the emphasis on the regional disparities of patenting activity and the identification of research clustering across the many European regions (chapter 3), an additional part of the study will be related to the analysis of inter-regional co-patenting linkages and thus to relational aspects of research and patenting activity (chapter 4). Therefore, the following empirical review discusses existing results, methodological issues and technical problems of alternative approaches. Finally, the review offers several arguments in favor of a pan-European clustering study and the identification of inter-regional co-patenting linkages.

2.2.2. Regional Disparities, Urbanization and Research Clustering

The majority of studies are primarily concerned with the effects of concentration and co-location but not with the general trends in the global distribution of research activities. It is generally assumed that the distribution of research activity is highly skewed because it was always skewed. However, researchers should pay more attention to measuring the spatial distribution and dynamics of research activity before empirically approaching the potential effects of specialization and/or concentration. What is more important? Knowing that skewed distributions may have certain effects or knowing that the distributions are skewed but that regional disparities/ concentration may show some meaningful dynamics? Accordingly, the analysis of regional disparities of innovative activity and research clustering seems to be of central importance. Furthermore, the analysis should combine an "industrial," "geographic" and "technological" dimension. Regarding this aspect, the empirical analyses in chapters 3 and 4 are primarily concerned with the distribution of patenting activity across the 819 European TL3 regions since the 1980s, the identification of research clusters and the structural analysis of inter-regional co-patenting linkages. Thus, the focus is on the distribution of knowledge-intensive tasks and the embeddedness of regions.

Regarding the theoretical base of concentration and disparity studies, it is necessary to distinguish four broad families of theoretical models that address the spatial distribution of research activity, as has been presented in the theoretical part of this study, especially in section 2.1: (i) the traditional neoclassical approaches predict convergence and dispersion; (ii) models with external scale effects, spillovers and different forms of externalities predict persistent disparities, depending on the initial distribution; (iii) models with internal scale economies and pecuniary externalities predict core-periphery structures or equity, depending on the initial distributions and the strength of centripetal and centrifugal forces; and finally (iv) models that include (ii) and (iii) differentiate between severe core-periphery structures and equity distribution, depending on the initial distributions and the strength of centripetal and centrifugal forces (Brülhart, 2001; Baldwin and Martin, 2004; Harris, 2008).[221]

From an empirical perspective, a strong motivation for analyzing the structure and dynamics of the distribution of European inventorship activity from patent data is based upon the fact that innovation data, and regional data in general, are said to show a strong and persistent non-normal distribution (i.e., skewness, kurtosis) and spatial autocorrelation (Fotheringham *et al.*, 2002; Scherngell, 2007; Anselin, 2007).

Several seminal studies already pointed out the non-normal, highly skewed distribution of gross domestic product (GDP) per capita, gross value added (GVA) and employment within and across European regions (Combes and Overman, 2004; Frenken and Hoekman, 2006; Paas and Schlitte, 2007, 2008). Moreover, a meaningful number of regional studies has highlighted stable agglomeration patterns in selected cities and urban regions or single countries. Kim (1995), for example, found that the correlation of the coefficient of regional localization for 2-digit industries at the US state level is around 0.64 between 1860 and 1987. Similarly, Dumais *et al.* (2002) identified stable agglomeration patterns at the 3-digit level for US-industries. With respect to plant location, Fujita and Ishii (1999) analyzed the location of R&D activities (mass-production vs. trial plants) and found that diversified trial plants are mainly located in central Japanese metropolitan areas, e.g., Tokyo, Kyoto. In opposition, plants for mass-production are mainly located in rural areas or in at least smaller cities, where urbanization externalities (and Jacobs externalities) are less prevalent.[222] Moreover, it has been demonstrated that the distribution of research (patenting) activity was already highly skewed in the past centuries. Pred (1966) examined US patent data for the mid-nineteenth century and found that innovative activity in cities was four times greater compared to the national average of patenting activity. Similarly, Higgs (1971) found that the number of US patents (between 1870 and 1920) was positively related to the urbanization level (see also Carlino, 2001). At the national level, Degner and Streb (2010) recently showed that foreign patenting activity in Germany was dominated by leading (and neighboring) countries, e.g., the United Kingdom and the United States, and that foreign patenting activity of countries varied with respect to the different waves of technological progress and diffusion of general purpose technologies.

Related to the recent decades, Puga (1999, 304) exposed that the European regions with the highest manufacturing employment density (27 NUTS regions) account for approximately

[221] The interested reader can take a closer look on the progress in the research lines (ii), (iii) and (iv) in the detailed literature review in chapter 2.1.

[222] Refer to Duranton and Puga (2001) for further interpretation.

50% of all manufacturing employment in the Union and for 45% of its population but only for 17% of the Union's areal surface. In comparison, the 14 US states with the highest manufacturing employment density also account for approximately 50% of the country's manufacturing employment, but only for 13% of areal surface and 21% of US population. Accordingly, employment and production seem to be more concentrated in the US. Regarding innovative activity, Audretsch and Feldman (1996) analyzed the spatial concentration of US production and innovation activities at the state-level. Their study was seminal in showing that innovation is, on average, more concentrated compared to production. In a European context, Maurseth and Verspagen (2002) argued that several European regions could be characterized as high-technology clusters, whereas Europe, i.e., the ERA, also includes low-technology dominated regions (especially in Southern Europe and the NMS). Similarly, Audretsch and Fritsch (2002) reported the existence of four different technological growth regimes for 74 Western German planning regions between 1983 and 1998.[223]

Several studies have made extensive use of patent statistics in order to analyze the dynamics of research and patenting activity in a spatial, sectoral and technological context. Concerning distributional particularities of innovative activity across European regions in the 1990s, especially in the ERA, in particular Caniëls (1996, 1997, 2000) and Breschi (2000) represent seminal contributions. The authors emphasized the highly skewed distribution of patenting activity, although at a very aggregated regional level (NUTS1/2). Paci and Usai (2000b) similarly contributed to the agglomeration of innovation and production debate, although they focused on a smaller set of countries and larger spatial aggregates.[224] Asheim and Gertler (2005) concluded that knowledge, and innovative activity in general, are geographically clustered, and that the tendencies towards geographic concentration have become more distinct over time. Florida (2005) also argued that the world, and the geography of knowledge, is "spiky," which reinforces the stylized fact that knowledge bases and research activity seem to be remarkably concentrated in space. Moreno *et al.* (2005a) found that specialized European innovation clusters, i.e., European NUTS1/2 regions, exist and that specialization is still increasing. Their empirical result is in direct opposition to common findings of studies on production clusters, which normally argue that production is continuously undergoing geographic dispersion.[225] Similarly, Paci and Usai (2009) used a population of 175 NUTS regions and demonstrated that EPO patent citations show strong core-periphery structures, with a deep gap between core-regions in central-northern Europe and the European periphery. These findings will be challenged by the empirical analysis of the geographic distribution of EPO patenting activity in chapter 3. Related to this debate, Castellacci and Archibugi (2008) argued that the explanation of the variance in the distribution of knowledge among 131 nation states can be reduced to two major factors: (i) differences in technological infrastructures and human skills, i.e., absorptive capacity, and (ii) differences in the creation and diffusion of codified knowledge.[226] The latter point will be empirically challenged in chapter 4.

223 Audretsch and Keilbach (2004) similarly explored the effects of entrepreneurship capital for German regions.

224 These studies only explored European patenting at a very aggregated regional (and national) level.

225 In this respect, Usai (2008) argued that one possible interpretation for concentration is that firms' strategic innovation activities may be to a large extent influenced by spatially localized interactions.

226 See Malecki (2010) for a similar point of view.

However, a severe issue of empirical investigations of core-periphery structures and spatial dynamics is the question of aggregation from a sectoral and a spatial perspective. Geographical economics and economic geography always have the problem of defining and defending the relevant industrial and geographical scale of analysis (Amiti, 1999; Ciccone, 2002; Arbia and Petrarca, 2010). Ideally, real-world industries and regions correspond to their theoretical counterparts (Brakman *et al.*, 2005). In practice, there is a trade-off between industrial detail and regional detail. Some researchers choose three-digit manufacturing industries, which are available at the national level (NUTS0/TL1). Other researchers instead prefer one-digit industries, which are also available at a more detailed regional level. The geographical scope of the new economic geography literature is, according to Brakman *et al.* (2005), by and large restricted to regional levels of analysis (NUTS2, NUTS3, TL2 and TL3 level). Brakman *et al.* (2005, 7) also suggested that *"[...] there is something to gain from sacrificing even more industrial detail for the sake of regional detail."* The methodologies and empirical analyses applied throughout this study directly follow their line of reasoning.[227]

Unfortunately, the availability of spatially disaggregated European data is disillusioning. Combes and Overman (2004, 2847), among others, recently complained that

> "[a]fter reviewing the literature, and given our first hand knowledge, the only conclusion that we are able to reach is that the European data are a mess. It is not clear where blame for this situation lies. It is clear that part of the problem stems from the institutional framework within which most EU governmental statistical agencies work. In particular, the fact that they often have no mandate to facilitate the re-use of data collected to fulfill their institution roles. Even where they do have a mandate, data are often expensive and incentives to ensure efficient delivery appear to be limited. [...] To summarise, the data situation is not good at the national, regional, or urban levels in the EU, although individual countries may provide excellent data sources."[228]

Accordingly, regional studies have to deal with several statistical issues, e.g., incomplete data coverage, selection biases, small sample size or inconsistent spatial and sectoral classification (Combes and Overman, 2004; Brülhart and Traeger, 2005; Scherngell, 2007). Consequently, most studies which center regional disparities of production, knowledge-intensive industries or research and patenting activity are either conceptualized at the

227 On account of this, this study explores 43 technology fields based on EPO patent applications by priority date. Additionally, the study explores structural dynamics of 6 large high-technology fields (laser, aviation, computers and automated business equipment, micro-organism and genetic engineering, communication technology and semiconductors). The patent classification system (IPC)-technology field concordance is applied to an own relational EPO patent database at the very disaggregate spatial level of OECD Territorial Level 3 (TL3) regions, which explicitly approaches the issue of functional spatial units.

228 For an identical critique refer to Arbia *et al.* (2005).

national level[229], or they are mainly organized at the level of large regional aggregates[230], or restricted to single countries[231] The trans-regional structures and dynamics of clustering remained unexplored in most studies, especially the distribution and clustering of knowledge-intensive tasks, i.e. patenting activity. Regarding pan-European and worldwide patent statistics, the OECD and EUROSTAT generally provide comprehensive statistics, reports and overviews (OECD, 2007a, 2008, 2009a). Nevertheless, the officially available reports are determined by a significant lack of disaggregated technology field calculations, a lack of disaggregated spatial classification systems and heterogenous statistical methodologies. Moreover, the reports are in most cases restricted to short periods.

To conclude, to the author's knowledge, no single pan-European empirical study exists that analyzes the spatial distribution of research and patenting activity at the regional TL3 level for all 819 regions of the EU-25, Norway and Switzerland for a comprehensive number of technology fields or industries for the last three decades. Regarding this deficit, the empirical analysis of the distribution of patenting activity in Europe in chapter 3 is a first essential objective of this study and aims to help to sharpen the cognition and to enrich the understanding of spatial structures, regional disparities and ongoing dynamics of research and patenting activities in Europe. In this respect, the empirical analysis will be dedicated to analyzing whether technology fields in Europe show decreasing disparities within the last three decades. Moreover, a harmonized, multidimensional research clustering index will be introduced, which represents the base for global cluster statistics and for the identification of leading innovative places in Europe.

2.2.3. The Regional Knowledge Production Function

2.2.3.1. The Origins of the Knowledge Production Function

As has been demonstrated in the last sections of the chapter, the distribution of innovative activity represents a central issue. Obviously, the analysis of potential effects of research clustering on innovative capacity and output (patents) represents a second step. The knowledge production function (KPF) represents a pivotal empirical approach which combines the "industrial," "geographic" and "technological" dimension of agglomeration economies and regional knowledge production. In applying the KPF approach, the main exercises aimed at measuring the spatial scope of spillovers.

229 See, e.g., Archibugi and Pianta (1992), Amiti (1999), Brülhart (2001), Midelfart-Knarvik *et al.* (2003), Midelfart-Knarvik *et al.* (2004), Aiginger and Pfaffermayr (2004), Greif and Schmiedl (2006), Legler and Krawczyk (2006). In addition, the majority of contributions at the level of European regions center GDP and GVA distribution, employment and unemployment dynamics (Frenken and Hoekman, 2006; Paas and Schlitte, 2007; Brakman and van Marrewijk, 2008). The distribution of US R&D labs is analyzed by, e.g., Carlino *et al.* (2010).

230 See, e.g., Combes and Overman (2004), Scherngell (2007), Rodríguez-Pose and Fratesi (2007), Brakman and van Marrewijk (2008), Brakman *et al.* (2009), Paci and Usai (2009).

231 See, e.g., Maurel and Sédillot (1999), Keilbach (2000), Greif (2001), Dekle (2002), Litzenberger and Sternberg (2006), Fingleton *et al.* (2007), Dewhurst and McCann (2007), Breschi (2008), Fornahl and Brenner (2009). Exceptions are Frenken and Hoekman (2006) and Paas and Schlitte (2007). Combes and Overman (2004) give a comprehensive overview regarding shortcomings of existing regional studies.

The knowledge production function is based on the work of Griliches (1979) and Griliches and Pakes (1980a), among others.[232] A first issue with regard to the KPF is the choice of the level of aggregation: the plant-level/firm-level, the sector- and industry-level, the regional level (functional and administrative units). Audretsch (1998), Foray (2004) and Malecki (2010), among others, described the KPF approach as a "comfortable world" of standardized models, in which only some agents, institutions and sectors are included in the production of knowledge. The KPF approach is considered to end up with a "black box" of knowledge production and diffusion (section 2.1.7).[233]

Following the original contribution of Griliches (1979), the KPF can be written (and finally estimated) in logarithmic form as in equation 2.2.1:

$$Q_i = \alpha + \beta_1 C_i + \beta_2 L_i + \beta_3 RD_i + \beta_4 Z + \varepsilon_i \tag{2.2.1}$$

Firm i's output (Q_i) is linked to the traditional inputs capital (C_i) and labor (L_i), but also to internal R&D activities (RD_i), and additionally to spatial spillovers (Z) that exhibit an effect on firm's output. Such externalities can originate from spatial, technological (sectoral/industrial) and/or social proximity. ε_i finally represents a random disturbance term and α is a constant term.

In the KPF tradition, the firm level is perhaps one of the most explored levels of research (Audretsch and Feldman, 2004).[234] Early production function approaches have been mostly dedicated to the individual level (firm- or plant-level) (Griliches, 1979; Griliches and Pakes, 1980b; Griliches *et al.*, 1984). Jaffe (1986) is one of the most cited studies, who has applied technological distance via patent data to the firm-level and regions. The approach is based upon the idea that industries foremost apply intra-industry knowledge, meaning that there is an implementation and absorption of technologically related knowledge.[235] Differing regional levels of aggregation lead to varying point estimates (and differing inference and causalities). This issue is well related to the modifiable area unit problem (MAUP) due to aggregation and zoning (Arbia and Petrarca, 2010). The spatial effects of knowledge spillovers, either intra- or inter-industry, or intra- or inter-regional, can vary largely with the spatial level.[236] It is worth noting that, according to several studies, the KPF model seems to hold for regions but becomes less compelling at the level of the firm or plant (Bergman and Usai, 2009; Malecki, 2010). Finally, a notable advantage of the KPF approach is that

---

232 Refer to Griliches and Pakes (1980b), Griliches *et al.* (1984), Hall *et al.* (1986), Griliches (1992b), Griliches (1990), Audretsch and Feldman (1996), Acs *et al.* (1997), Audretsch (1998), Feldman (1999), Keilbach (2000), Porter and Stern (2000), Feldman (2000), OECD (2009a), and Feldman and Kogler (2010) for further information and an overview.

233 A comprehensive overview is given by Feldman (1999), Scherngell (2007), Bergman and Maier (2009), OECD (2009a), Foray and Lissoni (2010), Feldman and Kogler (2010).

234 According to Griliches ([1992] 1998, 252), the empirical background is related to innovation externalities as *"[t]he more difficult to measure and the possibly more interesting and pervasive aspect of R&D externalities is the impact of the discovered ideas [...] on the productivity of the research endeavour of others."* See also Jaffe (1986), Feldman (1994a), Audretsch and Feldman (1996) and Acs *et al.* (1997).

235 Additional firm-level studies are contributed by Mansfield (1986). For an overview refer to Scherer (2005).

236 KPF analysis at the disaggregated micro level of plants/firms, establishments, or even lines of business, however, render the model of the knowledge production function less compelling.

it can be conceptualized at the agent-level, the plant-/firm-level, the city-level, but also the level of districts, counties, regions or countries.[237]

2.2.3.2. The Regional Knowledge Production Function

Several authors have modified the original knowledge production function by focusing on regions, which represents the "spatial dimension" of R&D clustering and agglomeration economies (Jaffe, 1989; Acs *et al.*, 1997; Acs, 2002).[238] It is a standard approach to apply formal R&D data at the spatial level. Unfortunately, such data largely ignore the complex processes of technological diffusion and knowledge accumulation, whereby tacit knowledge (section 2.1.7) is built up. This approach also overlooks, e.g., formal and informal institutions, history (Audretsch and Feldman, 2004; Malecki, 2010).

The basic knowledge production function (equation 2.2.2) is modeled including an innovative output ($INV_{i,t}$) and a vector of inputs, whereas the most important one is R&D (i.e., R&D expenditures or employment, $RD_{i,t}$). The KPF can, in general, be considered to be an unrestricted Cobb-Douglas production function (Verspagen, 1993; Usai, 2008).[239]

The regional knowledge production function has been heavily criticized as it models only global spatial processes (global mean regression) without the explicit recognition of region-specific (internal) set-ups. Moreover, the approach is simplifying the microeconomic issues (foundation) of the knowledge diffusion process, as it can only incorporate puristic regional spillover effects, i.e., global spatial processes. However, the pivotal advantage of the production function is its application to a large number of spatial units by means of small costs of modifying the econometric design (Audretsch and Feldman, 2004). Additionally, opposed to KPF regressions, survey-based research methodologies can hardly be generalized (Acs, 2002; Usai, 2008; Malecki, 2010).

$$INV_{i,t} = RD_{i,t}^{\alpha} e_{i,t} \tag{2.2.2}$$

Innovative output ($INV_{i,t}$) can be approximated by product introductions, new established processes or, more general, by patent applications or granted patents of different patent offices (e.g., EPO, JPO, USPTO, WIPO).[240] Issues are caught by the stochastic error term

237 It has been verified only for large macro areas (regions, countries). However, the KPF model becomes less convincing at the level of the firm due to the weak relationship found between R&D inputs and innovative output. In this respect, the application of GIS based data and distance models seems to add complexity and efficiency to econometric research (Audretsch and Feldman, 2004). Note, however, that one essential issue of spatial (dis-)aggregation is the modifiable area unit problem, which essentially influences the raise and decline of modeled spatial effects (and inference).

238 It was mainly Jaffe (1986, 1989) who contributed with insights related to university research spillovers on private firms. See also Keilbach (2000), Usai (2008) and Freund (2008) for an overview.

239 For a detailed discussion and additional information refer to Griliches and Pakes (1980a), Jaffe (1989), Coe and Helpman (1995), Audretsch and Feldman (1996), Acs *et al.* (1997), Audretsch and Feldman (1999), Bottazzi and Peri (2000), Anselin (2000), Varga (2000), Acs *et al.* (2002), Bottazzi and Peri (2003), Greunz (2003b), Greunz (2003a), Greunz (2004), Greunz (2005), Moreno *et al.* (2005c), Moreno *et al.* (2005a), LeSage *et al.* (2007), Scherngell *et al.* (2007), Crescenzi *et al.* (2007b), Crescenzi and Rodríguez-Pose (2008), Usai (2008), OECD (2009a), Ponds *et al.* (2010).

240 The empirical work in this project is highly abundant on EPO patent applications, because it is broadly accepted in empirical studies that applications seem to be stronger linked to the time of

($e_{i,t}$), which controls for unobserved determinants and random shocks.[241] The implementation of additional factors should grant higher efficiency and add some explanatory power to the production function.

Regional differences in R&D elasticities are interpreted as region-specific effects stemming from unique innovation infrastructure, which can also reflect private and public sector R&D characteristics (Freund, 2008; Usai, 2008). Therefore, R&D-activities can be disaggregated to their sectors of origin; i.e., the business sector ($\text{BusinessRD}_{i,t}$), governmental entities ($\text{GovRD}_{i,t}$), and the higher-education sector ($\text{UnivRD}_{i,t}$). Equation 2.2.3 represents such a production function.

$$INV_{i,t} = BusinessRD_{i,t}^{\alpha_1} GovRD_{i,t}^{\alpha_2} UnivRD^{\alpha_3} e_{i,t} \qquad (2.2.3)$$

Regarding spatial interaction and spillovers, several studies have criticized the sole implementation of region-specific inputs. Therefore, the inclusion of external factors seems logical as external factors and spatial spillovers may represent significant determinants of regional innovative activity and regional development (Crescenzi *et al.*, 2007b; Usai, 2008).[242]

Although the knowledge production function approach is determined by some methodological issues, there is a wide consensus that the positive relationship between R&D input and patent output is significant and strong at the regional level, and that spatial interdependence represents an essential phenomenon with significant impact on patenting activity (Audretsch and Feldman, 1996; Usai, 2008).[243] According to the aforementioned points, it can be argued that a non-normal distribution of R&D activity should explain the skewed distribution of patenting activity. Thus, the analysis of the geographic distribution of research and patenting activity seems to be crucial. In the following, selected studies on the US-American and European case are briefly summarized.

2.2.3.3. Knowledge Flows and R&D Spillovers in Europe and the US

Regional knowledge production is considered to exhibit significant inter-regional spillover effects. In this regard, studies applying the KPF approach combine the "industrial," "technological" and "geographic" dimension of knowledge production (see sections 2.1.6.5 and 2.2.1). In the following, selected studies at the regional level are reported with special focus on distance decay effects of R&D spillovers (i.e., spatial dependence).[244]

invention than granted patents. Moreover, most studies see statistical evidence for a time lag between two and three years (Greunz, 2005; Frietsch and Schmoch, 2006; Fraunhofer, 2009).

241 The latter variable is extensively challenged in early production function and convergence studies (e.g., the Solow residual in growth literature).

242 Oerlemans *et al.* (2001, 347), among others, concluded that *"[u]nder the condition of low problem levels, innovator firms utilise relatively more internal resources to innovate successfully. [...] In the case of highly complex innovation processes, this inwardness is no longer possible. The number and nature of innovation problems force innovators to utilise external resources."*

243 For additional conclusions refer to Funke and Niebuhr (2000a), Funke and Niebuhr (2000b), Niebuhr (2000), Paci and Pigliaru (2001), Greunz (2003a), Greunz (2004) and Freund (2008).

244 For a more detailed overview refer to table B.1 in the appendix. See also Keilbach (2000), Acs (2002), Usai (2008) and Freund (2008) for an overview.

Several studies have reported empirical evidence that the structure of knowledge production, the inter-regional transfer and diffusion of knowledge and the location of research activity may differ between Europe and the United States (Crescenzi *et al.*, 2007b; Usai, 2008; Kroll, 2009). In the following, selected knowledge production function studies on the United States and Europe are presented. An additional summary is reported in the appendix.[245] For the European case, the scarcity (and heterogeneity) of data at the regional level and several statistical reforms have prevented the formation of a consensus and major stylized facts. The number of regional studies is still relatively small compared to studies at the national level.[246]

In a US study, Acs *et al.* (1997) analyzed the impact of university and private R&D activities on regional innovation output with a knowledge production function. The authors used data from high-technology US-firms. The analysis was executed by applying cross-sectional data for states and metropolitan statistical areas (MSA) in order to estimate a knowledge production function with industry and university R&D as covariates. They found a positive impact of industrial and university R&D on local innovative output. More precisely, they depicted a positive and significant impact of university R&D on innovative activity with rather limited spatial range (approximately 50 miles around the university). However, they could not report a significant positive result for private R&D activities. They did not find statistical evidence for a significant and positive contribution of private business R&D to university research.[247]

Acs *et al.* (2002) analyzed the US MSA level.[248] Similar to Acs *et al.* (1997), they concluded that university research has a spillover range of approximately 50 miles around MSAs. Spillovers from private R&D, however, show tendencies to be contained within regional borders without inter-regional significant effects (negative but not significant), although elasticities for intra-regional business R&D are five times higher compared to university R&D (compared to state level regression). For addressing model misspecifications by means of spatial dependence, they additionally introduced spatial lags of the intra-regional covariates (business R&D, university R&D) by means of concentric rings (threshold distance of 50 and 75 miles), which may support the assumption of strong distance decay effects beyond MSA borders. The ML estimation (auto-regressive model) showed approximately the same results as the OLS method with distance decay of spatial lags (cross-regressive).[249]

245 For a summary, refer to table B.1 in the appendix. There exists an innumerable quantity of country-level studies. For further details see Verspagen (1997), Usai (2008), Freund (2008) and OECD (2009a). Verspagen (1993), among others, estimated a patent production function in a pooled cross-country time series data set. The sample consists of the 24 OECD countries and 13 newly industrialized countries (NIC). The elasticity of R&D activity is well above unity pointing to increasing returns to R&D intensity. Note, however, that the author only studies the country-level, meaning that inter-regional variation is not addressed.

246 The subsequent review places emphasis on selected European studies (published between 2000 and 2009). Nevertheless, the review is non-exhaustive.

247 These results support the hypothesis that different incentives may exist within the public and private sector, which may rather promote new knowledge to circulate in the public and university sector.

248 For similar results and a critical review refer to Keilbach (2000), Acs (2002) and Usai (2008).

249 The overall regression fit of their knowledge production function specification, which includes R&D indicators is above 0.5 ($0.599 \leq R^2 \leq 0.661$). Respectively, the regression fit of their extended knowledge production function model is above 0.7 ($0.718 \leq R^2 \leq 0.763$).

The authors reported no evidence that private R&D is endogenous to university R&D in the MSA.[250]

Varga (2000) contributed with a study at the level of US states and the MSA, showing that knowledge spillovers do not only exist within metropolitan areas. Varga reported evidence for significant positive spillovers from neighboring metropolitan areas up to 75 miles which has been similarly reported by Acs (2002).[251] Finally, a spatial range of 50-75 miles showed robust result for geographical externalities (knowledge spillovers) in the US MSA case.

Bottazzi and Peri (2000) studied European regions. In calculating the number of patents per square kilometer of all regions (large regional units) under analysis as the dependent variable (patent density), the authors tested the effect of regions' R&D expenditures (per square kilometer) and the overall influence of R&D expenditures within predefined distance bands (concentric rings), i.e., 0-300, 300-600, 600-900, 900-1,300 and 1,300-2,000 kilometers away from the regional center. For spatial spillovers resulting from R&D expenditures (cross-regressive) and patent applications (autoregressive), a significant positive impact on innovative activities in neighboring regions was reported. However, the spatial covariates only showed up with a significant positive sign for a distance band up to 300-600 kilometers, which represents a much larger distance, compared to the 50-75 miles distance bands reported in US studies.

Bottazzi and Peri (2003) similarly measured the extent of localized knowledge spillovers for 86 large European NUTS regions. Distance bands were identical to Bottazzi and Peri (2000). However, the analysis was complemented by controlling for technological distance. They defined a 30x1 vector (technology proximity indicator) for each region out of 625 IPC fields. Technological proximity was tested by generating the correlation coefficients of the technology vectors between the 86 regions.[252] Their spatial lag of regional R&D expenditures was significant for a 0-300 kilometer distance band. When disaggregating the 300 kilometer distance band, only R&D expenditures up to 100-200 kilometer showed a significant positive effect. The used patent data solely represent a 1/100 random extraction of EPO patent applications (6010 patents in total).

Greunz (2003a) estimated knowledge production functions for 153 European NUTS1/2 regions and reported a significant and positive effect of R&D expenditures, pursued in the first-, second- and third-order contiguity-based neighboring regions (median distances 91, 176, and 248 miles), on regional patenting activity. The spillover effects were not significant anymore beyond 250 kilometers. Technological distance was additionally included into the model by means of 118 IPC fields (technology sections), which turned out to show a significant impact of technologically lagged R&D controls.[253] The efficiency of the model increased with additional technologically lagged controls. Moreover, she offered the interesting result that national borders matter significantly in European regions in terms of patenting and that spatial significant R&D spillovers are mostly mediated by the business sector.

250 In opposition, they argued that there is evidence that university research in an MSA is endogenous to private sector R&D activity.

251 The model specification is performed in an OLS and IV set-up, which yields an overall regression fit above 0.599 ($0.599 \leq R^2 \leq 0.781$).

252 The overall fit of the spatial model is above 0.70 ($0.70 \leq R^2 \leq 0.91$).

253 The overall fit of the spatial model is above 0.9 ($0.92 \leq R^2 \leq 0.93$).

Lim (2004) used the knowledge production function approach to estimate the effect of specialization, diversity and competition on patents per capita at US-MSA level, which covered 313 observations. The author applied different estimators (OLS, ML, 2SLS, robust OLS). Specialization (localization) and diversity (urbanization) were both positive and significant, whereas only the spatially lagged diversity control was significant and positive.[254] However, the model set-up did not include R&D controls.

In a similar set-up, Moreno *et al.* (2005c) estimated a knowledge production function for 138 NUTS1/2 regions with spatial lags of 0-250, 250-500 and 500-750 kilometers distance. The authors reported significant positive effects from the first two spatial contiguity-based R&D expenditure lags.[255] The regression was done in an OLS- and ML-environment. The test statistics for remaining spatial dependence, i.e., LM-ERR and LM-LAG, were not significant due to several included control variables. In most cases, regional GDP (gross domestic product per capita) and the manufacturing employment share showed significant and positive point estimates, although spatial lags of business sector R&D expenditures seemed to decrease the significance of manufacturing employment.[256]

Bilbao-Osorio and Rodríguez-Pose (2004) estimated a knowledge production function for 103 European NUTS1/2 regions in a cross-sectional set-up. They reported significant positive effects on patenting activity from regional GDP and business R&D expenditures.[257] Opposed to peripheral regions, European regions were not affected by university R&D. Moreover, the regional stock of patents, the patent growth rate and the size of the high-tech sector exhibited significant and positive effects on GDP growth for the whole sample of European regions.

Moreno *et al.* (2005a) estimated a knowledge production function for European NUTS1/2 regions (similar to Greunz (2003a)). R&D activities of first-, second-, and third-order neighboring units had a significant and positive effect on the knowledge output of the spatial unit. Spatially lagged R&D activities showed strong distance decay effects, meaning that spillovers occur at a proximate distance. R&D expenditures at the first- and second-order distance contributed with elasticities around 0.22, whereas the effect from third-order neighbors was quite smaller (0.17). R&D expenditures at the 4th-order contiguity level were not significant. Moreover, Moreno *et al.* (2005b) estimated a knowledge production function for 175 NUTS1/2 regions with contiguity based spatial lagged controls, addressing

[254] The model fit is around 0.4 ($0.448 \leq R^2 \leq 0.461$).

[255] It is worth noting that several contributions to the European case operate at a very aggregated level, which offers a rather small sample of cross-sectional data; mostly at the NUTS1 level.

[256] The overall model regression fit was above 0.9 ($0.908 \leq R^2 \leq 0.918$) for both cross-regressive alternatives. The fit of the auto-regressive model was above 0.899 ($0.899 \leq R^2 \leq 0.908$). Moreno *et al.* (2003) similarly estimated a knowledge production function for 138 regions (respectively 123) with spatial lags of patent applications per capita as dependent variable. Additionally, they estimated a version with spatial lags of R&D expenditures (contiguity based). The overall model performance was above 0.9 ($0.908 \leq R^2 \leq 0.915$) for the latter version, whereas the former version had an R^2 above 0.899 ($0.899 \leq R^2 \leq 0.908$). All equations of the basic knowledge production function were estimated in an OLS framework (spatial cross-regressive model); LM-ERR and LM-LAG were not significant due to spatially lagged control variables.

[257] However, in most cases, GDP and R&D expenditures are highly correlated with patenting potentialities/patent applications. In this respect, GDP and R&D expenditures are highly correlated, which eventually introduces a bias. The overall model fit is similar to the above discussed contributions ($0.74 \leq R^2 \leq 0.85$), although their work does not contain spatially lagged variables.

auto-regressive interdependence.[258] The knowledge production functions were estimated in an OLS- and ML-framework for different sectors/industries. Additionally, the knowledge production function was complemented with technological distance controls that showed up with significant coefficients.

Greunz (2005) reported no direct inter-regional effect from governmental R&D expenditures. However, it is worth noting, that the results showed that the business and university R&D sector were both positively affected by lagged governmental R&D expenditures. Opposed to Acs *et al.* (1997), who have not found evidence for spatial dependence of private business sector R&D but only for university R&D, the results of Greunz (2005) indicate that inter-regional knowledge spillovers mainly originate from business sector R&D activities.

OhUallachain and Leslie (2007) estimated a knowledge production function for 50 US states. They showed that commercial patenting is highly dependent on R&D expenditures, whereas business R&D is positive and highly significant; university R&D and governmental/federal R&D is insignificant or has a significant negative effect on commercial patenting.[259]

Crescenzi *et al.* (2007b) estimated knowledge production functions for US patent growth for the period 1990-2002.[260] Their estimations for the US case covered 266 MSAs (145 MSAs/CMSAs respectively). They demonstrated that knowledge spillovers in the US do not cross a 80-110 km distance band, which supports the hypothesis of highly localized spillovers. The authors argued that the United States show significant research clustering and strong distance decay effects.[261] In comparing Europe and the US, Crescenzi *et al.* (2007a) reported a much higher population density in European regions compared to the US, indicating that major European metropolitan areas are located at a proximate distance. They argued that this could be one reason for a much stronger circulation of knowledge in Europe, which is reflected by stronger spatial autocorrelation of patenting activity. In this respect, the authors observed much stronger significant and positive effects on annual patent growth rates for 96 European NUTS1/2 regions compared to US MSAs. Spatially lagged R&D expenditures for the European case were significant although the authors introduced several additional controls (e.g., country dummy variables, agglomeration indicators, industry specialization, social filter). Interestingly, population density was not significant for Europe but significant and positive for the 266 US MSAs. The Krugman index, which measures specialization of regional employment, was significant and negative for European regions.[262]

Hauser *et al.* (2008) criticized existing (and recent) regional knowledge production function regressions by means of model misspecifications. The authors argued that the incorporation of social filters (e.g., political interest, friendship ties, trust, associational activity and technology and self improvement), generated by factor analysis, minimizes nuisance spatial

[258] The overall model performance is represented in a regression fit above 0.43 ($0.43 \leq R^2 \leq 0.86$).
[259] The regression fit is above 0.7 ($0.725 \leq R^2 \leq 0.870$).
[260] Earlier knowledge production function contributions used the stock of patent applications (applications/grants per year).
[261] The regression fit of the models are above 0.12 ($0.12 \leq R^2 \leq 0.32$).
[262] The regression fit for European regions was around 0.3 ($0.21 \leq R^2 \leq 0.47$).

dependence.[263] Their regressions are based upon 51 European NUTS1 regions. EPO patent applications per million inhabitants (log) were used as the dependent variable. R&D activity (aggregate of business, government, university, non-government) showed a significant and positive effect on patenting activity. This effect was six times larger compared to skilled employment in high technology sectors (HRST).[264]

Finally, in an OECD context, Usai (2008) estimated a knowledge (patent) production function by using PCT applications (Patent Corporation Treaty). The regressions covered 30 OECD countries and 61/271 spatial units (271 TL2 OECD regions, 61 North American regions, 201 European regions). The econometric results showed a significant and positive effect from intra-regional R&D activity and neighboring regions' R&D spillovers. Similarly, the spatially lagged dependent variable in an auto-regressive model was significant and positive.[265] However, the applied spatial lags in the cross-regressive setup seemed to differ from earlier contributions as the first-order contiguity lag of R&D expenditures was not significant after implementing the second-order contiguity-based lagged R&D covariate. Moreover, population density was not significant. It can be concluded that the latter two results are endogenous to the aggregation level as the OECD TL2 classification solely consists of large (macro) areas. This issue also applies to all of the aforementioned studies that were conceptualized at the aggregated NUTS1/2 level.[266]

To conclude, almost all reviewed KPF studies have demonstrated that intra-regional R&D activity and patenting activity are strongly correlated, that R&D activity exhibits a significant and positive effect on regional patent output, i.e. patent applications or granted patents, that inter-regional R&D spillovers are significant and positive but undergo strong distance decay effects, and that the distribution of research and patenting activity, although measured in most studies at a very aggregated spatial level, is (highly) skewed. Nevertheless, the origins (and micro-foundations) of regional spillovers remained a "black box" in the presented studies and the interpretation of such spillovers as knowledge externalities is certainly misleading (Breschi and Lissoni, 2001a; Breschi *et al.*, 2005).[267] Furthermore, regarding the quantity of existing KPF studies (at the NUTS1/2 level) and the remaining poor data availability, the estimations of regional KPFs seem to have hit fairly decreasing returns. Unfortunately, harmonized R&D statistics below the NUTS2 level do not exist, which represents another meaningful reason to approach the presented research questions with an alternative methodology.

263 Hauser *et al.* (2008, 869) concluded that *"[t]he spatial concentration of social capital is as important as the concentration of R&D and human capital in explaining observed autocorrelation of innovation."*

264 The overall regression fit of their setup was above 0.8 ($0.87 \leq R^2 \leq 0.90$).

265 The overall model efficiency of the European cross-regressive knowledge production function (spatial lag of R&D) is $R^2 = 0.908$, compared to $R^2 = 0.683$ for North America (United States) and $R^2 = 0.897$ for the whole OECD sample.

266 In opposition, the empirical analyses in this thesis focus on TL3 regions.

267 For additional national studies refer to, e.g., van der Panne (2004), Autant-Bernard and Massard (2007), Fritsch and Slavtchev (2007b), Richter and Freund (2008), Freund (2008), Arancegui *et al.* (2008), Andersson and Gråsjö (2009) and Patuelli *et al.* (2010). See also table B.1, appendix, for an overview.

2.2.4. Localization, Urbanization and Regional Development

As has been demonstrated in the last section, regional knowledge production is associated with strong regional spillovers at a proximate distance. Another prevalent debate in geographical economics and economic geography, i.e., the "industry dimension", centers the effects of regional industrial structures on regional employment growth, productivity, innovative capacities and innovation output. With respect to the latter, the debate is also known as MAR-Jacobs externality debate as theoretically discussed in section 2.1.6.[268] It can be argued that this debate primarily centers the "industrial" dimension (see also table B.1, appendix, for an overview).

Glaeser *et al.* (1992), among others, concluded, after having analyzed the top five industries and employment structures in US cities, that a higher diversity is associated with higher growth rates. Industry specialization, on the other hand, reduces urban employment growth.[269] In comparison, Henderson *et al.* (1995) argued that specialization appears to matter more for mature industries and technology fields. Conversely, urban diversity is essential for establishing new industries, which links the debate to the life-cycle concept (of regions, industries and clusters). Moreover, according to the theoretical review of urbanization economies and innovation externalities, the generation of new knowledge and the production of new products and services shows tendencies to be more concentrated in metropolitan areas which show a diversified industry structure.

Recently, researchers also developed interest in analyzing the effects of spatial technology and industry structures on research and innovative activity. Duranton and Puga (1999, 8) concluded that

> "[n]ot only is the creation of new plants biased towards larger and more diverse cities, but so is the location of innovative activities that lead to new products."

Similarly, Audretsch and Feldman (1999) argued that cities are the places where innovation occurs, and focused on the effects of Jacobs externalities from local industry structure. In using data on US product innovations they concluded that more than 90% of innovations are generated in metropolitan areas but that these spatial units account for approximately 30% of the US population. Furthermore, they found that regional industry specialization has a negative effect on innovative output, whereas city size and diversity across industries with a common science base have a significant and positive effect. However, Audretsch and Feldman (1999) could not find any positive effect from localization (specialization) on employment growth and innovative activity. Their results give support to the idea that highly localized industries, profiting from static and dynamic localization economies, may represent mature industries or clusters that rely on large scale production and intra-industry knowledge transfer (see also Duranton and Puga, 1999; Feldman, 2000; Audretsch *et al.*, 2008).

[268] The observed studies differ in their econometric methodologies and techniques and the dimension and specificity of the used database. The heterogeneity of the empirical results is mirrored to a certain degree in the heterogeneity of empirical approaches.

[269] Glaeser *et al.* (1992) analyzed the growth of industries in 170 US cities (1956-1987) in order to find empirical evidence for specialization and/or diversity. The authors did not find statistical evidence for MAR, but positive significant coefficients for diversity on industry growth. It is also reported from other studies that doubling city size increases productivity by 3-8%.

Strongly related to the "industrial" dimension of agglomeration economies, the so-called "specialization-diversity" debate was enriched by contributions that focus on "technological relatedness" and "related variety" and the effects of industry structures on knowledge transfer, productivity and employment growth (Neffke *et al.*, 2009; Boschma and Frenken, 2009a).[270] Jaffe *et al.* (1993), for example, found out that knowledge spillovers are not confined to closely related technologies (or industries); the authors argued that around 40% of patent citations did not come from the same patent class as the originating patent (refer to section 2.2.5). Accordingly, relatedness can be seen as an essential aspect of industry dynamics. Frenken *et al.* (2007) concluded that regions with a higher degree of variety among related industries will be determined by stronger local knowledge spillovers. Boschma and Frenken (2009a) also pointed to the importance of technological relatedness, arguing that new industries can connect to existing industries via various channels of knowledge transfer due to cross-fertilization. Similarly, Boschma and Iammarino (2009) brought forward the argument that regions may benefit from other regions via inter- and intra-industry knowledge flows as already theoretically discussed in section 2.1.7.5.

Although there exist many empirical studies on the "specialization-diversity" debate in an industry and city context, empirical evidence on the effects of the local industry structure (mostly employment) seems at best inconclusive (de Groot *et al.*, 2009; Beaudry and Schiffauerova, 2009). Feldman (2000) conceded that clear-cut answers remain elusive as long as empirical findings on urbanization and localization tend to vary.[271] de Groot *et al.* (2009) reviewed regression coefficients from more than 30 empirical studies in a meta-study. Both types of agglomeration economies, localization and urbanization, showed positive coefficients as often as they did the opposite.[272] After classifying and reviewing different sources of agglomeration economies Duranton and Puga (1999, 24) similarly concluded that

> "[this] does not imply that one type of city is economically more desirable than the other. [...] Some cities specialise in churning new ideas and new products (which requires a diversified base [...]), whereas other cities specialise in more standardised production (which, in turn, is better carried out in a more specialised environment)."

According to their understanding, the different types of local economic environment may matter at different stages of a product's (and industry) life-cycle (Duranton and Puga, 1999, 2001).

To conclude, almost all regional knowledge production function studies confirmed the existence of distance decay effects of knowledge (R&D) spillovers, irrespective of their origin

[270] Refer to Boschma and Iammarino (2009) and Neffke *et al.* (2011) for an overview.

[271] Refer to Audretsch and Feldman (1996), Audretsch and Feldman (1999) or van der Panne (2004) for further discussions.

[272] Glaeser (2000, 92) concluded that *"[...] for the moment, the role of concentration [i.e., of localization] and diversity does not seem to have been resolved by the literature. Different time periods and different samples give different results which suggests that there is no universal truth on this topic."* Similarly, Breschi and Lissoni (2001a, 5) have argued that *"[...] all the best-known studies on localised knowledge spillovers (LKS) seem to be unanimous in concluding that knowledge spillovers, either intra-industry or inter-industry, are important and strongly bounded in space."* Beaudry and Schiffauerova (2009, 334) argue that *"[the] analysis of the evidence presented in the paper strongly hints at measurement (level of aggregation of both industrial and geographical classifications) and to some extent at methodological (MAR and Jacobs indicators) issues as the main causes for the divergence observed in the literature and to the fact that the debate regarding MAR or Jacobs externalities remains unresolved."*

(R&D activity, patents).[273] Nevertheless, central issues have to be mentioned: (i) the regional classification represents a problem as aggregation from small to large units induces an averaging process which induces and/or enforces spatial autocorrelation between observations (Arbia and Petrarca, 2010); (ii) fractionally counted patent data automatically induce some kind of spatial autocorrelation in empirical analysis if a significant fraction of patent application originates from research activities with co-assignees/ co-inventors from neighboring regions; (iii) evidence regarding inter- and intra-industry effects remains inconclusive. Accordingly, the issue of regional "diversity vs. specialization" will be challenged by an alternative methodology (see chapter 3, section 3.5).

2.2.5. Patent Citations, Paper Trails and Real Spillovers

An alternative strand of research, which is very popular in studies on research clustering, R&D spillovers and knowledge diffusion, is the patent citation approach. It can be argued that the citation approach is the answer to elementary critiques regarding the existence, importance and micro-foundations of knowledge spillovers (Krugman, 1992, 2011). The approach combines the "industrial," "technological" and "geographic" dimension of knowledge production.

Adherents try to directly measure the extent of knowledge flows by using patent citation data. The analysis is characterized by the attempt to reconstruct paths of knowledge diffusion, i.e., paths of citations included in patent documents and their specific location and distance (Verspagen and Schoenmakers, 2000; Fischer *et al.*, 2005; Scherngell, 2007).[274] The citation approach allows ex post statements about the spatial range of knowledge spillovers and flows. In applying this approach, the research methodology represents an attempt to follow a "paper trail" that is left by citations (Feldman, 2000). Jaffe *et al.* (1993, 578) challenged Krugman's famous neglect regarding knowledge diffusion and argued that

> "[knowledge spillovers] do sometimes leave a paper trail, in the form of citations in patents."

An advantage of the citation analysis is that citation data can be applied in detail on specific technologies (IPC sections) and agents, which enables sector- and/or technology-specific conclusions. However, citations in patent documents are in most cases added by professional patent examiners at the patent offices (EPO, WIPO or USPTO) but not exclusively by the inventor and/or applicant (Alcacer and Gittelman, 2004; Scherngell, 2007; Criscuolo and Verspagen, 2008). Accordingly, included citations may not reflect the stock of knowledge of the person at that time when the patent was applied for; it may, in the other extreme, rather represent the detailed stock of knowledge of the patent examiner. Criscuolo and Verspagen (2008) showed that the share of patents with all citations included by the inventor has been constantly declining (from 10% in 1985 to 5% in 2000), while the fraction of patents with all citations added by the examiner has been rather constant. Additionally, they showed that the shares of all citations added by

273 For a final overview refer to table B.1 in the appendix.

274 See also Jaffe *et al.* (1993), Maurseth and Verspagen (1999), Keilbach (2000) and Maurseth and Verspagen (2002).

EPO examiners instead of inventors differ tremendously. In organic chemistry, e.g., almost 15% (65%) of all citations are added by the inventor (examiner), while in information technology only 2% of all citations are added by the inventor (93% by examiner). Their results also support the importance of spatial distance for EPO patent citations (Criscuolo and Verspagen, 2008).

The patent citation approach does neither capture knowledge flows via co-inventorship activity, i.e., co-patenting linkages, nor knowledge transmission via researcher mobility in networks, labor migration (brain drain, brain gain), knowledge flows between firms and between customers and firms (Scherngell, 2007; Fischer *et al.*, 2009; Paci and Usai, 2009).

The perhaps most prominent contribution in patent citation analysis is Jaffe *et al.* (1993) (and Jaffe and Trajtenberg (2002)), who used a "case-control-matching approach" in order to verify patent citations as a potential transfer channel of knowledge spillovers. The authors compared the location of cited patents with the location of citing patents by using the official inventor location.[275] As Jaffe *et al.* (1993, 579) have argued:

> "[W]hy should innovations tend to cluster spatially more in some industries than in other industries [...]? The most difficult problem confronted by the effort to test for spillover localization is the difficulty of separating spillovers from correlations that may be due to a preexisting pattern of geographic concentration of technology related activities. That is, if a large fraction of citations to Stanford patents comes from the Silicon Valley, we would like to attribute this to localization of spillovers. A slightly different interpretation is that a lot of Stanford patents relate to semiconductors, and a disproportionate fraction of the people interested in semiconductors happen to be in the Silicon Valley, suggesting that we would observe localization of citations even if proximity offers no advantage in receiving spillovers. Of course, the ability to receive spillovers is probably one reason for this pre-existing concentration of activity."

In their study, Jaffe *et al.* (1993) calculated two probabilities: (i) a patent cites another patent registered by a nearby agent (e.g., university, firm), with both patents referring to the same technology and having originated from a similar point in time; (ii) two patents are similarly geographically linked, without existence of formal links through such patent citations. The overall result of their analysis was that intra-national and intra-regional citations happen more often than one would expect from the distribution of patenting activity. Additionally, they observed that citations happen more frequently when the citing and cited patent belong to the same spatial unit (see also Keilbach, 2000; Scherngell, 2007). They found that knowledge spillovers are not confined to closely related industries or technologies. In this respect, they suggested that around 40% of patent citations do not come from the same patent class as the originating patent (Jaffe *et al.*, 1993; Keilbach, 2000). This result can be interpreted as an indication in the direction of Jacobs externalities and recent contributions to related variety and technological relatedness (section 2.1.6.3) (Boschma and Frenken, 2009b; Neffke *et al.*, 2009). Other studies, e.g., Malerba *et al.* (2003, 3) came to the conclusion that

> "[patent citations] can be regarded as a noisy signal for spillovers."

[275] See Fischer *et al.* (2005) and Scherngell (2007) for an overview.

Similarly, Alcacer and Gittelman (2004, 14) discussed the citation approach and concluded that

> "[o]verall, our results do not change the presumption that patents trace out knowledge flows: inventors face strong legal pressures to reveal all they know, and our results do show that inventor citations follow a pattern we would associate with inventor knowledge. [...] the bimodal pattern does not contradict that knowledge spillovers are localized."

Another study is the one of Maurseth and Verspagen (2002). They analyzed the impact of language and national borders on the knowledge diffusion across 112 spatial units in 14 countries. They concluded that there exists a negative correlation between flows of knowledge and spatial distance. They showed that knowledge diffuses more easily across countries when they have the same language.

Thompson and Fox-Kean (2005), in opposition, who built on the work of Jaffe *et al.* (1993), suggested that the localization of spillovers measured by patent citations was generally overestimated. Additionally, the authors critically discussed earlier studies. Similar criticism has been presented by Agrawal *et al.* (2003), who argued that a careful extraction of control patents (control samples) is a necessity for patent citation analysis as citations will automatically be co-located even in the case of absence of knowledge spillovers as soon as technology fields are geographically clustered (see chapter 3, section 3). It has also been questioned whether the results of citation analysis pertain to a perhaps too high aggregation level. Otherwise, it may be possible that earlier patent co-operations explain most of these spillovers (Thompson and Fox-Kean, 2005).

In a recent study, Scherngell (2007) analyzed patent citations of high-technology EPO patent applications between large European regions (NUTS1). He showed that geographic distance had the smallest negative effect in the electronic industry, whereas the pharmaceutical industry and aviation industry showed strong decay effects. Moreover, language and national borders had the expected negative impact on patent citations. Summarized, Scherngell showed that high-tech patent citations suffer tremendously from spatial distance and that the citation structure shows strong concentration and core-periphery structures. Additionally, he observed inter-regional patent citation linkages between leading European core regions. Peripheral regions, in opposition, are generally characterized by small numbers of received and made citations in almost all analyzed technology fields (see also Fischer *et al.*, 2009). Although Scherngell did not explicitly examine the regional typology of the regions under observation, urban regions tend to receive the largest fraction of European patent citations in the sample (e.g., Ile-de-France, Oberbayern, Stuttgart, Noord-Brabant, Darmstadt, Düsseldorf, Lombardia, Köln, Stockholm, Rhone-Alpes). This result is in line with the findings on research clustering and patent densities as will be presented and discussed in chapter 3 (section 3.5). Moreover, his results on citation linkages are similar to the results on co-patenting linkages between regions in this study (see chapter 4).

In a similar set-up, Paci and Usai (2009) offered an EPO patent citation analysis for a group of 175 NUTS0, 1 and 2 regions for 17 countries. Although the authors made use of an IPC-technology field concordance table for descriptives, they did not offer additional (relational network) results with respect to different technology fields. They concluded from their analysis that (i) citation links decrease with spatial distance, (ii) citation flows are higher between contiguous regions (that are sharing a common border), and (iii) citations happen more frequently between regions that have a similar technology base. Interestingly, the authors additionally argued that spatial distance has generally lost influence on patent citation intensity, which means that patent citations became less sensitive with respect to physical distance of researchers.[276] Paci and Usai (2009, 675) argued that

> "[t]his picture seems to indicate an increase in the spatial scope of knowledge diffusion which goes in the direction proposed in the literature under the label of "death of distance" [Cairncross, 1997]."

A similar development for scientific European co-publications was reported by Hoekman *et al.* (2010) from their gravity model estimation at the NUTS1/2 level (see section 2.2.7). Their study, although different with respect to methodology and aggregation level, points into the same direction as the results reported in this thesis, i.e., the inter-regional co-patenting network analysis (see chapter 4, sections 4.3.4 and 4.3.5).

Sonn and Storper (2008) argued that the localization of patent citations (i.e., the proportion of local citations) has increased within the last two decades, which can be interpreted as a significant increase in localization and concentration of knowledge spillovers. Their findings are in line with those of Paci and Usai (2009), although the latter argued that national borders became less important in the course of time.

Finally, Bergman and Usai (2009) reported that knowledge flows within the EU, measured via patent citations, are strongly localized in European core member states and that these flows emerge from a small number of strongly agglomerated places.[277]

The following conclusions can be drawn from the reviewed studies. Patents and their cited-citing ratio are highly concentrated in space. Spatial distance is said to hamper research collaboration intensity but negative effects from national borders seem to have vanished. For more details, the interested reader is referred to the above mentioned literature for further information.[278] However, although based upon relational data, the studies of Scherngell (2007) and Paci and Usai (2009) are problematic by technical reason. Most citation studies applied the standard NUTS classification, which may lead to a severe bias in network data (data on citation, co-patenting, co-publishing, among others). This is a crucial concern as the underlying spatial classification system shows a bias in the absolute number and size of regions included in the analysis, e.g., Denmark as a single region vs. 40 German regions (Paci and Usai, 2009).[279] For this reason, another spatial classification system is applied in the own empirical analyses in chapters 3, 4 and 5. Thus, potential risks that might originate from a problematic spatial classification system are prevented. Furthermore, the

[276] Paci and Usai (2009) also make use of an IPC-technology field concordance table; they apply the "Yale-concordance" and the one proposed by Schmoch *et al.* (2003). Due to the fact that the results have not changed, their basic scenario is based on the Yale-technology concordance. This is interpreted as another sign that the Schmoch et al. concordance is generally considered an established concordance table.

[277] For similar conclusions see Fischer *et al.* (2005), Scherngell (2007) and Fischer *et al.* (2009).

[278] Additionally, Feldman (1999), Audretsch and Feldman (1999), Feldman (2000), Breschi and Lissoni (2003), and Döring and Schnellenbach (2006) contributed with seminal overviews and discussions of the related literature.

[279] Other studies in this respect are Scherngell (2007), Hoekman et al. (2009), Maggioni and Uberti (2009), Hoekman et al. (2010).

citation approach is problematic, because it is not sure that knowledge spillovers, by means of documented patent citations, have really been realized. Almost 90% of all citations are traced by patent examiners, which raises severe doubts regarding the citation approach (Criscuolo and Verspagen, 2008). Moreover, the citation approach completely ignores the major fraction of knowledge that is frequently transmitted via the market process and within intra- and inter-regional network linkages, i.e., co-inventor networks (Ejermo and Karlsson, 2004; Maggioni and Uberti, 2009; Hoekman *et al.*, 2010). Accordingly, patent citations cannot be interpreted as a valid measure of (direct) interaction between individuals or regions.

Due to the presented methodological drawbacks and the ongoing dispute with respect to the patent citation approach, it will not be applied in the following empirical analyses. Regarding the mentioned deficits, it is also important to acknowledge that a parallel line of analysis exists, measuring knowledge flows by application of data on co-inventor networks. The co-inventor approach is reviewed in section 2.2.7. A detailed European co-inventor network analysis is favored, i.e., an EPO co-patenting network analysis at the regional level (see chapter 4).[280]

2.2.6. Researcher Mobility, Social Networks and Diaspora

Regarding the micro-foundation of knowledge transmission, another interesting working channel represents the interaction between persons in social networks that allows the non-codified transmission of "tacit knowledge" (see chapter 2, section 2.1.7.2). This approach is intertwined with studies on knowledge diffusion via labor markets, spatially mobile networks of researchers and diaspora networks (see chapter 2, section 2.1.7.5).[281] Moreover, the approach is partially intertwined with patent citation analysis. Accordingly, the social network-mobility approach combines the "industrial," "technological," "geographic" and "social" dimension of knowledge production, whereas the main focus is on the latter.[282]

Researchers are regarded as carriers of highly specialized, implicit knowledge. Related to this approach, a good starting point are the concerns of Breschi and Lissoni (2001a, 976), who argued that

> "[t]he role of geographical distance in the economics of knowledge transmission [...] is still rather controversial."

Therefore, it is of great necessity to analyze the spatial context and channels of knowledge transmission in more detail.

Almeida and Kogut (1999) and Zucker *et al.* (1998), among others, assumed that the spatial concentration of knowledge flows is related to the features of labor markets for highly skilled workers. There exist several studies that have examined how labor mobility of

---

280 A detailed European co-patenting network study at the national and regional level follows in chapter 4.

281 A discussion is offered in Burger et al. (2009), Bergman (2009), Ter Wal and Boschma (2009) and Franz (2010).

282 See also Almeida and Kogut (1999) and Trippl (2009).

inventors, and researchers and their networks act as a key mechanism for knowledge transmission (diffusion), which gave rise to research on social networks and diaspora (Almeida, 1996; Almeida and Kogut, 1999; Saxenian, 2006, 2007). Although localized knowledge spillovers rely on local networks and thus on geographical space, the local nature of knowledge transfer is explicitly based on the tacit nature of knowledge due to technology-specific determinants of inventorship (see section 2.1.7.5) but not primarily due to pure distance decay effects (see section 2.1.6.7). The network approach is related to the idea that informal knowledge exchange happens within social networks of inventors and their research collaborations (Gordon and McCann, 2000; Breschi and Lissoni, 2003). Breschi and Lissoni (2006) analyzed the structure of Italian inventor networks with special focus on the mobility of scientists. Breschi and Lissoni (2006, 9) concluded that

> "[i]t remains true, however, that many social networks dedicated to the production of knowledge as a club good are geographically bounded, since spatial proximity may help the network members to communicate more effectively and patrol each other's behaviour."

With respect to co-patenting activity of researchers, Breschi and Lissoni (2009, 439) recently argued that

> "[t]he most fundamental reason why geography matters in constraining the diffusion of knowledge is that mobile researchers are not likely to relocate in space, so that their co-invention network is also localized."

Breschi and Lissoni (2006, 8) furthermore center club good characteristics of such networks and the features of epistemic communities that determine the knowledge transmission process within (and between) communities:

> "[S]pillovers from an active club member will reach distant fellow members with some delay or imprecision, and will possibly never reach outsiders. [...] To the extent that many [social] networks are concentrated in space, co-localisation would appear as a significant determinant of access to spillovers."

Moreover, with respect to patent citations, Breschi and Lissoni (2004, 14) argued that the spatial dimension (see section 2.2.5) is strongly related to the structure and dynamics of social networks. They concluded that

> "[t]he population of inventors is more than a tiny and unchecked sample of all individuals who can influence inventors themselves. Rather, it may possibly represent the most immediate and influential social environment from which inventors draw ideas and information, at least from technical contents of their patents."

Thus, these processes of localized labor mobility and informal knowledge exchange seem to be very sensitive to the underlying network structures (Burger *et al.*, 2009; Bergman, 2009).[283] When scientists and researchers co-locate, there is a high probability that most informal contacts between researchers also take place at a proximate distance. It is argued that face-to-face interactions offer the possibility of complex and intense forms of communication and interaction (von Hippel, 1994; Lissoni, 2001; Hoekman *et al.*, 2010). Inventor

[283] See also Almeida and Kogut (1999), Breschi and Lissoni (2001a), Breschi and Lissoni (2003), Breschi and Lissoni (2006), Agrawal *et al.* (2006), Breschi and Lissoni (2009) and TerWal and Boschma (2009).

networks generally rely on face-to-face contacts, which automatically gives a tacit nature to it (see section 2.1.7.2). The same argument was picked up by Breschi and Lissoni (2006, 8) who argued that

> "[knowledge spillovers] would be localized if and only if a significant proportion of social networks are also localized in space. [...] If those people move away from where they originally learnt, researched, and delivered their inventions, knowledge will diffuse in space. [...] That is, knowledge flows (where pure spillovers or traded services) are localised to the extent that labour mobility also is."

Concerning network developments, a first approach toward this hypothesis is to measure the effects of job-hopping in a spatial context. Therefore, local labor markets may represent superior levels of analysis opposed to larger administrative areas. In observing inter-regional labor flows, one may find a powerful explanation for spatial dependence in R&D and patent application activity; especially for explaining the geographical scope of research networks by technology class. Accordingly, some fraction of the investment in any innovation project will create technological externalities that positively affect other innovation projects as soon as researchers change their jobs or scientists exchange knowledge informally (Almeida and Kogut, 1999; Breschi and Lissoni, 2001a; Greunz, 2003a). Since labor mobility is considered being a regional phenomenon, knowledge spillovers and flows based on labor mobility are mostly localized (Balconi *et al.*, 2004; Breschi and Lissoni, 2009).[284]

In a US-study, Zucker *et al.* (1998) have analyzed the mobility of "star scientists" and their spatial range and underlying research networks.[285] They focused on the relevance of human capital for knowledge spillovers in a study of the geographical location of biotechnology firms. Their main research interest was restricted to the very early stages of innovations in this sector, where results of scientific research in universities or research institutions are transformed into commercial products. Their empirical study is based on the assumption that specialized knowledge is generally embodied in individuals. Zucker et al. showed that the geographical location of firms is closely related to the location of star researchers in the field. They also showed that the company, with which the scientist was in contact, used the new pieces of knowledge. Accordingly, the results of the authors can be interpreted as evidence that there are no (pure) knowledge spillovers, as the external knowledge of the researcher is implicit knowledge (i.e., embodied/tacit knowledge) (see also Döring and Schnellenbach, 2006; Breschi and Lissoni, 2009).

In another US study, Almeida (1996) focused on the relationship between spatial mobility of engineers in the US semiconductor industry and the localization of patent citations in this technology field. According to their reported results, there exists a considerable relationship between researchers' mobility and the spatial distributional structure of US patent citations. In a similar work, Almeida and Kogut (1999) have focused on the mobility patterns of patent holders (engineers) in different localized US industries. The authors

284 See also Almeida and Kogut (1999), Breschi and Lissoni (2001a), Breschi and Lissoni (2003) and Breschi and Lissoni (2004).

285 So-called "star scientists" are considered to represent the fraction of researchers, who are highly creative and productive and who discover "breakthrough technologies" (Feldman, 2000).

reported that the local transfer of knowledge across companies is endogenous to the inter-firm mobility of patent holders and that labor mobility is high but localized. They argued that Silicon Valley is one of the few clusters where mobility positively affects the innovative output of firms. The high mobility of researchers between firms in Silicon Valley is generally attributed to region-specific social institutions.

Singh (2005) referred to the network and citation approach in a US-study and found strong empirical evidence that social ties increase the probability of knowledge flows between individuals (measured by patent citations). In this respect, Singh combined the patent citation and social network approach. Identically to Breschi and Lissoni (2003), the author concluded that geography matters because interpersonal networks tend to be localized in a few places.

Oettl and Agrawal (2008) similarly argued that network linkages between researchers generally remain, although researchers frequently relocate in space (i.e., job hopping). They suggested that these linkages are reflected by patent citations by former colleagues. As the market does not (fully) price these flows, Oettl and Agrawal argued that flows of knowledge via labor mobility represent a kind of externality. Nevertheless, it can be argued that firms are generally aware of this external source of knowledge. Thus, experts and star scientists are in particular recruited because of their accumulated implicit knowledge (see also Bergman and Usai, 2009).

Breschi and Lissoni (2009) suggested that networking activity across agents and locations is responsible for a large fraction of localized knowledge flows between individuals (and regions). They concluded that the effect of non-market externalities (i.e., the spatial lag in knowledge production function models) was generally overestimated in past KPF studies due to methodological issues.

In a European context, Miguelez *et al.* (2009) recently used regionalized PCT patent data (EURO PCT) for studying the mobility of highly-skilled individuals as a possible mechanism of inter-regional knowledge transfer. Building on Breschi and Lissoni (2009), the authors hypothesized that knowledge flows are localized to the extent that inventors' mobility is also localized, which would explain the existence of local spatial dependence in explanatory spatial data analysis. Similarly to Miguelez *et al.* (2009), Miguelez and Moreno (2010) found strong support for the positive relationship between regional labor market mobility and regional patent densities for a sub-sample of European macro regions. They concluded that there exists a positive correlation between intra-regional labor mobility and regional patent applications.

To summarize, social network-mobility studies have, depending on the analyzed epistemic community, generally reported a highly localized mobility of researchers, which indicates that (implicit) knowledge transmission is localized to the extent that networks are localized. Furthermore, migratory movements of researchers seem to affect the "paper trail" of patent citations (see previous section) due to emerging and disappearing informal network linkages (i.e., social ties) and diaspora. This aspect has already been brought forward in the theoretical review (see section 2.1.7.5). Nevertheless, the statistical identification of researchers' mobility represents a meaningful issue in studies that intend to cover dozens of countries, hundreds of regions and many industries. With respect to this problem, a promising line of research is dedicated to the analysis of regional co-inventor networks, i.e., co-patenting networks, which will be discussed in the following.

2.2.7. Research Collaborations and Co-Patenting Networks

The internationalization of technology and R&D shows large cross-country differences (Guellec and van Pottelsberghe de la Potterie, 2001; Belitz *et al.*, 2006). Therefore, an alternative strand of research increasingly examines co-patenting structures in order to analyze the structures and dynamics of R&D collaboration activities in an international, regional and firm-level context (Maggioni *et al.*, 2007; TerWal and Boschma, 2009; Boschma and Frenken, 2010).[286]

Inventor/ co-patenting network analysis is said to have potentialities to contribute to the understanding of regional innovation systems and core-periphery patterns in knowledge intensive industries (Maggioni *et al.*, 2007; Burger *et al.*, 2009; Powell and Giannella, 2010). Besides the structural composition of linkages and networks (Kroll, 2009), recent empirical research places special emphasis on the changing structure of research networks (TerWal and Boschma, 2009; Burger *et al.*, 2009).[287] This approach combines the "industrial," "technological," "geographic" and "social" dimension of knowledge production, agglomeration economies and networks. In light of the previous theoretical discussion in chapter 2 (see sections 2.1.7.3, 2.1.7.4 and 2.1.7.5), inter-regional research networks are considered to represent pivotal factors that affect the geography of innovation.

It is also argued that co-patenting studies offer a way to directly measure international and inter-regional knowledge flows, i.e., an assessment of the globalization of applied R&D (Frietsch and Schmoch, 2006). Moreover, the analysis avoids several shortcomings and technical issues of the aforementioned approaches. Accordingly, the analysis of regional co-patenting networks can be regarded as a fruitful alternative (Johansson and Quigley, 2003; Ejermo and Karlsson, 2004; Iammarino and McCann, 2006).

The co-patenting approach is considered the only methodology that explicitly addresses the theoretical issues of inter-regional and extra-cluster research linkages in an appropriate way (see also sections 2.1.7.4 and 2.1.7.5). Moreover, unlike patent citations, co-patents detect the localization of researchers working on the same inventions. Therefore, co-inventorship can be regarded as a good approximation of intended technological and scientific collaboration (Maggioni *et al.*, 2007; Maggioni and Uberti, 2009). This has also been argued by Ejermo and Karlsson (2004, 2), who concluded that

> "[k]nowledge transfers should be qualitatively and quantitatively more substantial than citations as indicators of the overall flows of knowledge within an innovation system. After all, even if citations do reflect knowledge spillovers, deliberate cooperation must be of much larger magnitude than casual and random "spillovers." Co-authorship structures therefore seem more adequate for assessing the relative merits to the extent that knowledge travels across space."

[286] Burger *et al.* (2009) and Bergman (2009) represent comprehensive reviews of patent citation and network studies. See also Singh (2005), Lam (2007) and Lobo and Strumsky (2008).
[287] See also Orsenigo *et al.* (1997), Iammarino and McCann (2006) and Glückler (2007).

Co-inventor studies examine the structure and determinants of these collaborative patterns which are considered meaningful mechanisms of inter-regional R&D knowledge flows and spillovers (Ejermo and Karlsson, 2004; Bergman and Maier, 2009; TerWal and Boschma, 2009). Several co-inventor studies have been conceptualized at the national level (Ejermo and Karlsson, 2004; Ponds *et al.*, 2010). However, only a few contributions challenged the network structures beyond national borders (Kroll, 2009; Maggioni and Uberti, 2009; Hoekman *et al.*, 2010). Moreover, a detailed analysis of the distribution of these networks across geographic space is still missing in a pan-European context.

For the purpose of analyzing inter-regional networks of inventors the information on applicants and inventors in patent data is in particular valuable (Maggioni *et al.*, 2007; Paci and Usai, 2009; TerWal and Boschma, 2009). Accordingly, patent data are used as relational data.[288] Identical to Balconi *et al.* (2004), Ejermo and Karlsson (2006) and Maggioni *et al.* (2007), among others, the co-patenting analysis in this study (see chapter 4) implicitly assumes that researchers, who are listed within the patent application, know each other (personally) and share explicit and implicit knowledge in order to generate new technologies. Therefore, co-patenting activity is considered to be a meaningful proxy for the analysis of innovative collaboration activity between individuals and spatial units.[289]

Linkages between agents and regions can be ex post analyzed in order to identify co-patenting activity and co-inventor networks.[290] The major reason for taking the address of the inventor as the central selection criterion for localizing patents is that MNEs generally assign their patents to companys' headquarter locations. Accordingly, patents which are realized in firms' R&D subsidiaries will in most cases exhibit the address of the headquarter when using the applicant's address for analysis, although most of the inventors will be located as residents in the subsidiaries' regions (Verspagen and Duysters, 2004; TerWal and Boschma, 2009). Consequently, co-patenting studies allow researchers to distinguish between relatively open and integrated places (cities, regions, countries) and places that follow implicitly or explicitly a more closed (isolated) path of research activity.[291]

There exist several possible cases of co-patenting activity in a regional context: (i) research collaboration and co-patenting between local and neighboring units, (ii) co-patenting between local and foreign units, (iii) co-patenting between several local and foreign units.[292] In several technology fields international co-patenting, or at least regional border-crossing research activity, is driven by multinational corporations that vary considerably in their organizational structures. As R&D teams of a single company can be located across a number

288 In addition, co-patenting studies are extremely powerful when combined with covariates, which supports the interpretation of econometric results from co-patenting studies.

289 Refer to Wilhelmsson (2009).

290 This methodology, however, can only be applied in sectors and industries, in which intellectual property rights (i.e., patents) are generally used. Furthermore, this methodology is said to be biased towards (successful) inter-firm knowledge exchange and protection via patenting (TerWal and Boschma, 2009). For an analysis of co-patenting data the information on the inventor location can be applied to identify knowledge flows between agents, cities, regions and countries.

291 Furthermore, the data on R&D linkages can be enriched by data, if available, on foreign R&D labs in the country, foreign R&D expenditures or the number of foreign-owned inventions. However, studies also have to consider different sectors and technology fields because patenting intensities and co-patenting propensities vary considerably.

292 e.g., the agent-level, city-level, regional level, country-level.

of cities, regions or countries, it can be assumed that multinationals are responsible for a meaningful fraction of inter-regional R&D collaboration linkages (Belitz *et al.*, 2006; Lam, 2007; Fraunhofer, 2009), although large corporate R&D labs have fallen in prominence (Powell and Giannella, 2010). To conclude, it can generally be differentiated between two forms of inter-regional co-operation in R&D: (i) within-organization inter-regional collaboration due to joint research of inventors in different locations but affiliated within a single multinational company (cross-border knowledge flows without spillovers outside the company); (ii) between-organization inter-regional collaboration and joint research of inventors with different organizational and national affiliation who collaborate for an invention. The latter case is considered to induce knowledge spillovers beyond companies' borders.

A serious issue, however, is the judgement about the direction of inter-regional knowledge flows on the basis of co-patenting activity. It seems to be overly simplistic to assume that foreign research labs merely absorb knowledge from their host countries. It is argued that such laboratories are engaged in processes that involve knowledge exchange between domestic and foreign researchers (Fraunhofer, 2009). Moreover, co-patenting information only allows researchers to identify the residence (work place) of inventors, but not their nationality or path of migration.[293] As a consequence, statistical judgements can only be made about research collaboration intensities from co-patenting data, especially at the regional level. In this regard, Bergman and Usai (2009, 10) recently proposed that

> "[c]o-patenting is a good indicator of the localisation of inventors that have worked at the same invention and can be a good proxy for scientific and technological collaboration across space. [...] The starting point for a network analysis of the innovation process is the micro-level of individual agents. [...] by aggregating data at a given geographical level (i.e NUTS2 or NUTS3), one may build a geography-based innovator network."

With respect to the current state of research on co-patenting networks, TerWal and Boschma (2009, 742, 753) recently suggested that

> "[v]irtually no studies on the dynamics of the structure of networks in space exist [...]. [F]urther research is needed on how the structure of networks evolves over time and space and, particularly, how the evolution of networks is related to the evolution of clusters. [...] treating patent data as relational data provides us with considerable opportunities to study the dynamics of regional innovation networks, which is, till today, a rather unexplored though promising field of study."

Taking the aforementioned statements seriously, the empirical analysis in this study (see chapter 4) follows the methodological arguments and propositions of Bergman and Usai (2009) and TerWal and Boschma (2009) and places the emphasis on the structure of European research collaboration activities and co-patenting networks (see chapter 4, section 4.3.5).

In the following, results of selected co-patenting studies are briefly reviewed. The list of European studies remains however quite small (as the approach is relatively novel in literature) which again clarifies the need for additional empirical studies.

293 Accordingly, it would be definitely misleading to develop opinions about brain-gain or brain-drain between regions or countries based upon co-patenting data.

Research collaboration and knowledge co-production in co-inventor networks have been studied at the regional and national level. Andersson and Ejermo (2002) and Ejermo and Karlsson (2004) analyzed co-inventorship activity for Swedish regions based on patent data. Van Looy *et al.* (2003) analyzed co-patenting activity of knowledge generating organizations (e.g., institutes). In this respect, co-patenting between universities and public research institutes and industrial companies are becoming an increasingly important topic.[294] Moreover, institutional set-ups and practices vary tremendously (Martin *et al.*, 2008). The establishment of legal frameworks for university patenting also has an impact on co-patenting activity (Fraunhofer, 2009). Owen-Smith and Powell (2004) made use of a "pipeline concept" and discussed the transmission channels used in distant knowledge intensive interactions between agents. They concluded that physical distance is not the only important factor, even though knowledge flows (and spillovers) may be more effective within a regional network than across national borders. The authors showed that agents in the Boston biotechnology industry accessed knowledge via local interaction and strategic partnerships at an inter-regional and international distance. They argued that firms construct network linkages in order to benefit from research excellence abroad (see also Bathelt *et al.*, 2004; Powell and Giannella, 2010). However, regions need significant absorptive capacity in order to communicate and absorb forefront knowledge (Owen-Smith and Powell, 2004; Ejermo and Karlsson, 2006; Freund, 2008).[295]

In a Dutch study, Ponds *et al.* (2010) performed a network analysis for Dutch regions based on scientific publications. They came to the conclusion that physical distance essentially matters for scientific co-operation.

In a recent US study, Powell and Giannella (2010) analyzed the average spatial co-inventor distance (miles) by using USPTO patent data. They observed an increasing dispersion of co-patenting activity between the years 1975 (101-161 miles) and 2005 (215-185 miles) in the technology fields aerospace, biotechnology, optics, pharma/chemicals and semiconductors. Similarly, Johnson *et al.* (2006) reported that the distance between co-inventors has on average increased from 117 miles (1975) to approximately 200 miles (1999). They argued that emerging technology fields, e.g., computers, semiconductors and biotechnology, exhibit much stronger clustering than mature industries (and technology fields). They also argued that these technology fields have started to spatially spread within the last years. These findings are very similar to the computations for the European case reported in this thesis (see chapters 3 and 4).

In a European context, Maggioni *et al.* (2007) used co-patent data as one indicator in their analysis of the importance of traditional spatial spillovers vis-à-vis relational spillovers. They analyzed a sample of 109 European regions at the NUTS2 level within a gravity equation model. The authors combined data on the participation in the same research networks (EU Fifth Framework Programme) and EPO co-patent applications. In this way, they examined the factors that support patenting activity. The authors made the

294 Researchers in the public sector are increasingly managing their intellectual property. However, this issue is still pronounced differently in European member countries.

295 Related to these results and interpretations, Malecki concluded that *"[s]ome places are able to create, attract, and keep economic activity [...] because people in those places make connections with other places"* (Malecki, 2002, cited in Bathelt et al., 2002, 17). See also Cohen and Levinthal (1990) and Bathelt *et al.* (2004).

distinction between geographical and relational spillovers and structural features. They empirically tested if relationships, which are based on inter-regional networks between excellence centers, generally predominate research relationships at a proximate distance (contiguity). However, it is important to note that Maggioni *et al.* (2007) only analyzed a few European countries (France, Germany, Italy, Spain and the United Kingdom) and that they applied the very aggregated NUTS1/2 classification, which might implement several issues (e.g., spatial autocorrelation due to aggragation/averaging process).

In a later work, Maggioni and Uberti (2009) focused on international network linkages. Their analysis completely ignored intra-national linkages. The gravity model regressions covered data on internet hyperlinks, EPO co-patent applications, European networks of researchers and data on Erasmus student mobility. The authors concluded that knowledge linkages seem to concentrate in a few European NUTS2 "super-regions," which means that European structures are rather resembling "scale-free networks" but not "small worlds." Again, it has to be noted that the sample of regions included in their analysis, although at the well-known NUTS2 level, is heterogenous. This generally implies distorted network structures as the number of unique and overall linkages (edges) is endogenous to the regional classification system.[296]

Kroll and Mallig (2009) offered a comparison of US and European co-inventor network structures. Their main objective was to examine differences with respect to the spatial scope of network linkages. Their analysis is static as it does not offer a dynamic network comparison. Nevertheless, they clearly demonstrated that US networks are by and large more localized than European networks, although the applied European spatial classification system is rough (NUTS1 level).

Hoekman *et al.* (2009) discussed results of their European co-inventorship analysis with special focus on scientific (journal) co-publications (Web of Science) combined with EPO co-patenting data. They analyzed 1316 European NUTS3 regions and argued that the majority of co-publications seem to happen at a proximate distance between several central network nodes. Their study gives additional indication that European research excellence is by and large dominated by a small number of European NUTS regions. However, the applied spatial classification system in their study is considered to be highly problematic as German regions account for 439 of all 1316 EU-27 regions (33.15%), which obviously implements a central bias into the generation and analysis of relational data.[297]

Similarly, in a recent study, Hoekman *et al.* (2010) analyzed the effects of geographical distance and borders on the intensity of research collaboration at a higher spatial level, now across 313 European regions in 33 EU countries. Based upon co-publication data for the years 2000-2007, the authors found that the tendency towards collaborations with partners at a proximate distance did not decrease, while the bias towards collaboration within countries did decrease.[298] Hoekman et al. concluded that the observed decreasing effect of national borders may be an indication towards an ongoing integration process of regions into the ERA. However, innovative collaborations are still sensitive to physical

[296] This issue is also challenged in the empirical analysis in this thesis in chapter 4.
[297] This issue is picked up as another spatial classification system is used in the empirical analyses of this study (refer to chapters 3, 4 and 5).
[298] Their results are quite similar to the conclusions of Paci and Usai (2009).

distance. Finally, it should be noted that the high level of spatial aggregation in their study, i.e., NUTS1-2 regions, tends to eliminate variation in regional co-patenting activity and to enforce spatial autocorrelation.[299] In view of this, the NUTS 1/2 level has been heavily criticized, e.g., in a recent study of Paci and Usai (2009, 672), who argued that

> "[t]he NUTS2level, or higher, is commonly used in the regional analyses based on European data even though the phenomenon under examination [co-patenting] would deserve some attention at a more disaggregated territorial level."

To take this critique seriously, research clustering and co-patenting activity will be analyzed at a more disaggregated level in the following empirical analyses in this study (see chapters 3 and 4).

Having reviewed studies in the co-inventor/co-patenting network tradition, it can finally be concluded that there exists a relatively small body of contributions to inter-regional co-patenting and core-periphery structures of research collaboration in a European context. An in-depth analysis of the structural dynamics of technology-specific inter-regional co-patenting networks is still missing. It is generally argued that Europe is determined by a meaningful dispersion of patenting activity and a significant dispersion and expansion of co-patenting activity, although the empirical evidence is rather weak. Regarding this meaningful deficit, the empirical analysis in this study will emphasize the structures and dynamics of inter-regional co-inventor networks and research collaborations between European regions in chapter 4. Besides a detailed analysis of foreign co-inventor activity (number, shares) at the national level (section 4.3.4), the empirical analysis places emphasis on inter-regional co-inventor networks at the TL3, TL2 and TL1 levels and on 43 technology fields (section 4.3.5).

299 Studies that make use of detailed co-patenting data are Maggioni *et al.* (2007), Kroll (2009), Maggioni and Uberti (2009), Christ (2009), Ponds *et al.* (2010) and Miguelez and Moreno (2010).