Specialization Patterns in Trade and Technology

Thesis submitted in accordance with the requirements
for the degree of

DOCTOR OF PHILOSOPHY

by Maria Luisa Mancusi

Supervisor: Prof. John Sutton

The London School of Economics and Political Science
University of London
United Kingdom

February 2004

ACKNOWLEDGMENTS

I would like to thank all the people who influenced my work and contributed in different ways to the completion of this thesis.

I first wish to thank my supervisor, Prof. John Sutton, for his advice and guidance throughout this work: I have learnt from him more than I can show in this thesis.

Academic, administrative and IT support staff at the London School of Economics and at the Suntory and Toyota International Centres for Economics and Related Disciplines have always been helpful and have provided an enriching working environment from which I highly benefited.

I also thank my friends, who often had to prove very understanding. Among them very special thanks are due to Elisabetta Falcetti, whose contribution to this thesis has been decisive. I shared with her all the frustrations and the excitements of research in the past years and if I am finally concluding this work it is also because she always gave me good reasons for doing it.

I am highly indebted to my family and my husband's family. My parents and my brother are a continuous source of strength and my husband's family has always given me unconditional support, regardless of the numerous difficulties it has been going through.

Finally, my greatest gratitude goes to Carlo, my husband, because he has born the burden of this work as much as I did and, most importantly, because he put his trust in me and made me trust myself.

ABSTRACT

This thesis focuses on the hypothesis of increasing returns and path-dependence in technological development, which originates in the supposedly localised nature of external economies from knowledge creation, and looks at its implications in the innovation literature and in recent models of endogenous technological change, specialization and growth.

Chapter 1 provides an empirical assessment of the hypothesis of increasing returns and path-dependence. A *recontracting process* formalises the idea that in the presence of strong national externalities a country's pattern of technological specialization tends to polarise towards extreme values thus leading to the emergence of a bimodal distribution. This prediction is found to be at variance with the data. However, reinforcing effects appear to be at work in situations of strong disadvantage. This might be the effect of scarce past research experience limiting countries' ability to absorb external knowledge.

The analysis of Chapter 1 is further developed in Chapter 2, where differences across technological fields are accounted for, to allow for the possibility that only some of them may be subject to increasing returns.

Chapter 3 compares technology and trade specialization patterns for a group of advanced countries. The analysis shows that their relationship is weak: this weakens the case for self-reinforcing mechanisms in technological change leading to persistence in trade patterns.

Chapter 4 finds that the elasticity of innovation to international spillovers is positive and significant, thus suggesting they may be an important force leading to mobility in technology and trade specialization patterns. Absorptive capacity positively affects the elasticity to spillovers, but its effect depends on the position of the country with respect to the world technological frontier: the larger the gap of a country with the technological leaders, the lower is its ability to absorb and exploit external knowledge, but the larger is its potential to increase this ability.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The theory of technological accumulation, as developed in Cantwell (1989) maintains that even over long time periods countries will keep and even reinforce their relative technological position: established patterns of technological specialization will be rather stable and characterised by areas of clear technological strength and areas of clear technological weakness. The assertion rests on the view that technological change is a cumulative rather than a random process for two main reasons. The first reason is that the directions of technical changes are often defined by the state-of-the-art of the technology already in use: each firm's learning is local (in the technological space) and specific to the firm's own search history. The second reason is that the probability of technological advances by firms and countries is, among other things, a function of technological levels already achieved by them. In particular, countries' technological accumulation in individual sectors rests on complementarities and interdependencies that give rise to a structured set of technological externalities, which are highly specific to particular spatial and institutional contexts. These externalities reproduce or even increase over time (Dosi et al, 1990). Thus technological change is a self-reinforcing process, that is a process subject to increasing returns.

Self-reinforcing processes in economics (in particular, in the economics of technology) and the implications of increasing returns have been formalised and studied by Arthur in a series of papers (Arthur, 1994). He shows how, in the presence of positive feedback, the probability of further steps along the same path increases with each move down that path, because the relative benefits of such steps compared to once-possible options increase over time. As a consequence, actors have strong incentives to continue down a specific path once initial steps are taken in that direction. Sequencing is critical in that different earlier events (different sequences) may produce different outcomes: history matters. Processes that are unable to shake free of their history, are said to yield path-dependent outcomes.

Although arguments about technology have provided the most fertile ground for exploring the conditions leading to increasing returns, economists also have applied increasing returns arguments to economic change more broadly. A prominent development in recent discussions of economic growth has centred on endogenous growth theory (Aghion and Howitt, 1998). This strand of research argues that increasing returns associated with economic applications of knowledge help accounting for the puzzle of growth rates (notably in developed countries during the post-World War II period) far greater than what measured increases in inputs of capital and labour could explain. Unlike capital and labour, many aspects of knowledge are non-rival: their use in one firm does not prevent their use in another. The introduction of a new good may then give rise to positive technological external effects when the invention reveals new technical possibilities which competitors can incorporate into the next generation of their own products without paying a fee to the original inventor.

Positive technological external effects may also be the source of positive feedback and path-dependence in the dynamics of specialization in open economies. Indeed, arguments about increasing returns have gained wide acceptance in recent analyses of international trade. Researchers began by focusing on economic trends which appeared anomalous from the perspective of traditional trade theory. If comparative advantage results from natural features of different countries, one would expect most trade to occur between quite different countries, such as North-South trade of manufactured goods for raw materials. Most international trade, however, is North-North: developed economies trade primarily with other developed countries, including extensive exchanges within particular industries. Again, increasing returns may provide an explanation. Knowledge-intensive sectors will be prone to positive feedback. Countries that gain a lead in a particular field, for whatever reason, may consolidate that lead over time. The result is a high degree of specialization. Even countries with similar initial endowments develop divergent areas of economic strength. Comparative advantage is not simply given, it is often created through a sequence of events unfolding over time.

Within the literature on endogenous technological change, the reinforcement of initial specialization patterns results either from sector-specific learning-by-doing (e.g. Krugman, 1987) or from research and development (e.g. Grossman and Helpman, 1991, ch. 8) in the absence of international knowledge spillovers. Whether patterns of trade exhibit path-dependence and lock-in depends on the cross-border mobility of technology, which also determines whether there is convergence or persistent differences in levels of per capita output across countries. The importance of the issue has motivated a series of empirical studies on the scope of knowledge spillovers. The effect of international knowledge spillovers has been evaluated by assessing the impact of a pool of external knowledge on either total factor productivity (i.e. within a standard production function framework) or directly on innovative performance (i.e. in a knowledge production function framework). The pool of external knowledge is usually represented as the amount of R&D conducted elsewhere weighted by some measure of proximity in the technological or geographical space, taken to be representative of the intensity of knowledge flows between the source and the recipient of spillovers. Contributions within this strand of research widely differ on the way such knowledge flows are inferred. These differences are then reflected in an equally wide range of results.

Inferring flows of knowledge from flows of goods, Coe and Helpman (1995) find that international spillovers from foreign R&D positively affect productivity growth. Their methodology has been criticised because knowledge spillovers may be confounded with rent externalities and some authors have provided econometric evidence that casts doubt on the effectiveness of trade as a mechanism for knowledge transfer (Keller, 1998; Eaton and Kortum, 1996 and 1999). Other studies have inferred the intensity of knowledge flows from geographical distance (e.g. Bottazzi and Peri, 2003) or from the distance in the technological position of the source and destination countries (e.g. Branstetter, 2001). These studies mainly suggest that spillovers are more intra-national rather than international in scope. Finally, recent studies have employed patent citations as "paper trail" of knowledge flows, although the discussion has mostly been focused on the goodness of citations as a proxy of knowledge flows rather than on the impact of national vs. international spillovers when the intensity of knowledge flows are inferred from the intensity of citations (a notable exception is a recent paper by Peri, 2003). The findings on the scope of knowledge spillovers still remain inconclusive: while a border or geographical distance effect seems to emerge from some studies, others have found that this has been decreasing in time (Keller, 2000) or that the reach of knowledge flows depends on the sector/technology involved and on whether they originate from a technological leader (Peri, 2003). However, with only few very recent exceptions (Griffith, Redding and Van Reenen, 2001; Griffith, Harrison and

Van Reenen, 2003), have these studies investigated the extent to which the effectiveness of international knowledge flows depends on the recipient country's ability to understand and benefit from external knowledge.

This thesis contributes to the above mentioned strands of economic literature in a number of ways. Empirically, if the explanation of the stability in trade specialization patterns was indeed based on the reinforcing effects characterising technological progress (due to either sector-specific learning by doing or to localised knowledge externalities), persistence should be particularly pronounced in technological specialization, where learning and the positive external effects in the form of knowledge spillovers from R&D have their most direct and strongest impact. It is then interesting to study the empirical dynamics of technological specialization in industrial countries in order to verify whether it shows any of the implications of increasing returns. The evidence is shown to be at variance with these implications (chapter 1). In particular, technological specialization patterns appear to be characterised by high mobility. This is consistent with the empirical finding of substantial mobility in recent trade analyses (Proudman and Redding, 1998 and 2000) and calls for greater attention to the forces leading to mobility in specialization patterns, such as international knowledge spillovers. There are, however, two qualifications to these findings. First, high persistence is found in situations of strong disadvantage. A possible interpretation is that the inability of countries to move out of such situations is related to their scarce past research experience, which limits their ability to absorb external knowledge. Second, there are indications of differences across technological fields in the dynamics of technological specialization: in one case this dynamics shows features compatible with the existence of self-reinforcing mechanisms (chapter 2). Technology and trade specialization patterns are then directly compared for a group of advanced countries. Their relationship is weak and mobility in technological specialization is higher than in trade specialization, thus further weakening the case for self-reinforcing mechanisms in technological change leading to persistence in trade patterns (chapter 3).

Drawing on the previous findings, the last part of the thesis (chapter 4) then looks at the effectiveness of international spillovers in determining a country's innovative performance. The empirical analysis is performed on highly disaggregated data, accounting for differences across technological fields, and the elasticity of innovation to spillovers is modelled as a function of *absorptive capacity* (Cohen and Levinthal 1990). The elasticity of innovation to international spillovers is found positive and significant, thus suggesting they may be an important force leading to mobility in technology and trade specialization patterns. Furthermore, absorptive capacity positively affects the elasticity to spillovers. This provides a potential explanation to the observed dynamics of technological specialization and suggests the importance of incorporating into theoretical models the determinants of the ability to benefit from spillovers.

Looking at the contributions of each chapter in detail, chapter 1 provides an empirical assessment of the relevance of the hypothesis of path-dependence in the sectoral technological development of open economies, a hypothesis related to the supposedly localised nature of external economies from knowledge creation in the innovation literature and in recent models of endogenous technological change, specialization and growth.

The evolution of technological specialization is illustrated with reference to a *recontracting process*, which formalises the idea that in the presence of strong national externalities (when international spillovers are absent or weak) countries should display the tendency to specialize in selected technologies or, equivalently, patterns of technological specialization should display the tendency to polarise towards extreme values thus leading to the emergence of a bimodal distribution (i.e. one should observe increasing overall specialization in technologies). Established technological specialization profiles should then display persistence, with particularly low mobility between the extreme ends of specialization: the event of a technology moving from high (low) to low (high) specialization should be extremely unlikely. This tendency should be particularly pronounced for countries with already unbalanced initial specialization patterns.

In the empirical analysis I represent a country's pattern of international specialization in technologies through the distribution of the country's relative innovative output shares across technologies. Innovative output is measured using patent applications at the European Patent Office (EPO) and relative shares are obtained as a modified version of the Balassa (1965) index of revealed comparative advantage: the resulting measure of specialization is the Revealed Technological Advantage (*RTA*). The pattern of international technological specialization at any one point in time can be characterized by the distribution of *RTA* across technologies. Evaluating its dynamics over time requires an analysis of the evolution of the entire cross-section distribution. This involves two different, but related issues. On one hand, there is the issue of the changes in the overall degree of international technological specialization, which may be evaluated by analysing the evolution of the *external shape* of the *RTA* distribution. Do we observe an increasing specialization in a limited subset of technologies (a polarisation of the *RTA* distribution towards extreme values), or has the degree of specialization remained broadly unchanged? On the other hand, there is the issue of persistence versus mobility in international technological activities. This addresses questions related to the *intra-distribution dynamics*, such as: what is the probability that a technology moves from the upper (lower) quartile of the *RTA* distribution to the lower (upper) quartile?

The evolution of the *RTA* distribution is modelled adopting a distribution dynamics approach (Quah 1993a, 1996a, 1996b, 1996c), whose features make it particularly suited to study the two issues just described. The issue of changes in the external shape of the distribution is addressed by estimating the distribution of *RTA* across technologies in subsequent periods. Intra-distribution dynamics include information on switches in ranks and on the distance traversed when such switches happen. A way to quantify this phenomenon in the sequence of *RTA* distributions is to assume that the process governing the evolution of the specialization level of a country in a technology is a General Markov Chain. The law of motion of the sequence of distributions can then be described by a first order time-invariant autoregressive process, where the operator mapping the distribution from period t to period $t+k$ gives the conditional distribution of the specialization index at time $t+k$, given its value at time t. This operator is a Markov stochastic kernel or, when the continuous state space is partitioned into a finite number of intervals, a Transition Probability Matrix. Both can be estimated non-parametrically from the data and provide an interpretation of persistence as a measure of the probability that the *RTA* index of a country in a technology remains close to its initial value as time passes by.

The empirical dynamics of technological specialization emerging from the analysis of industrial countries does not seem to support the idea that there are cumulative and reinforcing mechanisms at work, which could then generate path-dependence in the original technology and, hence, trade specialization patterns. Countries do not show increasing specialization in a limited subset of technologies (a polarisation towards the extreme values of the distribution representing the specialization pattern), but rather the opposite. Technological specialization displays significant mobility: fluctuations around and far from initial levels happen with a probability almost always higher than 0.5. It is high specialization levels that display the lowest persistence as they tend to revert towards low levels. Both overall mobility and reversion from above are more pronounced for countries with higher overall degree of specialization.

These results are not in line with the core predictions of the theories of technological accumulation and path-dependence cited above, and cast doubt on the often alleged causality from hysteresis in technology to hysteresis in trade specialization. They suggest the importance of directly evaluating the effect of the forces capable of inducing changes in the international specialization over time, such as international knowledge spillovers. One feature of mobility emerging from the empirical analysis also deserves particular attention. Situations of initial complete or high de-specialization are characterised by relatively high persistence: it seems to be mostly difficult for a country to improve specialization in technologies where it is in a very disadvantaged position. This finding might signal that in the absence of a sufficiently high absorptive capacity, originated from previous experience in a technology, countries find it extremely difficult to overcome their weaknesses. Even if R&D spillovers were international in scope (a fact which could generate mobility in technology and trade specialization patterns), countries need to have some prior level of knowledge, R&D investment, or complementary assets in the relevant technology in order to understand and employ knowledge produced elsewhere.

The findings of chapter 1 call for the analysis of the relative importance of national vs. international knowledge spillovers and of the role of absorptive capacity in determining a country's innovative performance. This is done in chapter 4. Before that, however, two limitations of the empirical analysis of chapter 1 are considered. The evidence of substantial mobility in patterns of technological specialization is obtained without taking into account differences across technological fields. If only some technologies are subject to increasing returns (Arthur, 1994), then pooling the observations from all fields might obscure the signs of reinforcing effects in a subset of them. This issue is discussed in chapter 2. Finally, one of the reasons why the evidence of chapter 1 weakens the case for causality from hysteresis in technological development to hysteresis in trade specialization patterns is that mobility in technology appears higher than that emerging from analogous trade analyses (Proudman and Redding, 1998 and 2000). However, no direct comparison can be done because the analysis of chapter 1 is focused on technologies, which do not have a direct correspondence with industries or product groups. Chapter 3 takes care of this issue.

The aim of chapter 2 is to verify whether there exist technology specificities that might affect the evolution of a country's ability to innovate persistently in a particular field or to catch-up with actual leaders. Knowledge characteristics, cost structures, learning and externality effects may differ across technological fields and this may determine different technology dynamics, which might, or might not be characterised by reinforcing effects. In particular, the more the knowledge base in a technology field is complex, cumulative and firm-specific, the more one should expect a country in a relatively advantaged position to be able to reinforce it in the future and one with a relatively disadvantaged position to find it difficult to catch-up. The issue is then to see whether there is any field of technology where high de-specialization implies lock-in and, contrary to the generalised findings of chapter 1, high specialization induces positive reinforcing effects because of increasing returns in the creation of new knowledge.

The aim is achieved by studying the evolution of the cross-country distribution of revealed technological advantages in each of the following technological fields: electronics, instruments, chemicals and pharmaceuticals, processes, machinery, consumer goods and civil engineering. If technological specialization in any of these fields is characterised by reinforcing effects, then the distribution should again show a tendency towards polarisation: countries should display either high or low specialization in the field and their specialization level should persist over time. In trying to assess these tendencies, one immediate difficulty arises: the fields are quite widely defined. This is due to both features of the data and methodological constraints. As a consequence each field might comprise quite substantial heterogeneity because it includes high-tech, fast-growing segments together with old, stagnant segments. The dynamics of the two might be very different and such differences might be hidden in the aggregation. In spite of this limitation, the field of electronics shows some distinctive features that are consistent with the existence of reinforcing effects: countries tend to display either high or very low specialization in the field and to maintain their relative position over time.

Chapter 3 studies the relationship between a country's pattern of specialization in trade and its pattern of specialization in technology. Patterns of trade are determined not only by differences in technology and by technological change, but also by other factors, among which differences in relative factor endowments (Harrigan, 1997; Gustavsson et al, 1999). Therefore, one should expect, a priori, that the empirical similarities between the two patterns are limited. However, it is interesting to verify whether the size of the correlation between the two specialization patterns is affected by the aggregation level adopted. Empirical trade analyses are almost invariantly performed on highly aggregated industry data. If the correlation between technology and trade specialization patterns at this aggregated level is reasonably high, it will be valid to draw inferences from patterns of trade specialization as to underlying mechanisms such as path-dependence, which operates on technology. If instead the correlation is low, then such inferences are problematic. Chapter 3 then complements chapter 1 in that it directly compares mobility in technology and trade specialization: if any mechanism of path-dependence were at work, persistence in technological specialization should be at least as high as in trade specialization.

The analysis of chapter 3 is focused on the five most industrialised countries. As in the previous chapters, I use patents as a measure of innovative output and map their classification (the International Patent Classification, IPC) and the Standard International Trade Classification (SITC Rev. 3) into a common classification. Technology and trade specialization patterns and dynamics can be directly compared using this correspondence. The central theme emerging is that the correlation between technology and trade specialization patterns is extremely weak when using aggregated data, but positive and significant, although still low, when using disaggregated data. At such level of detail, technology and trade specialization levels show tendency towards convergence in the long run. Furthermore, persistence in technological specialization is significantly lower than persistence in trade specialization in the short-run, but becomes similar over a five-year horizon. The main implications are therefore that it is important to work with disaggregated data and to study the forces that lead to mobility in specialization patterns, in particular the role of international knowledge spillovers, emphasised by theory of dynamic comparative advantage and endogenous technological change.

The impact of knowledge spillovers is evaluated in chapter 4 within a knowledge production function framework using data on 135 micro-sectors in the chemicals, electronics and machinery industries for six major industrialised countries (US, Japan, Germany, France, UK and Italy) over the period 1981-1995. I assume that in country *h* firms operating in micro-sector *i* produce new knowledge using both their own R&D and external knowledge originated either elsewhere in the same country or in another country. To proxy new knowledge I use data on European patents. I then use patent citations to trace knowledge flows within and across countries among the 135 micro-sectors. Such flows are then used to obtain national and international knowledge spillover pools. Quite importantly, the national spillover pool is obtained using only citations to *other* national firms and institutions, hence excluding self-citations (i.e. citations to previous patents by the same applicant firm), which cannot be regarded as a "paper trail" of knowledge flows and which account for a large proportion of overall national citations.

Results from different empirical studies seem to suggest that knowledge spillovers are mainly intranational rather than international in scope (Jaffe et al, 1993; Branstetter, 2001; Maurseth and Verspagen, 2002). In one of these studies, Maurseth and Verspagen (2002) employ citations by patent applications at the European Patent Office (EPO) to trace knowledge flows across European regions: they find that patents are more likely to cite other national patents rather than foreign patents. In this chapter I show that this result arises because cross citations between European regions exclude all citations directed towards the world technological leaders (US and Japan). Once these are included in the analysis the home country effect disappears and the share of international citations is found to be particularly high in countries below the technological frontier. Consistently, international spillovers are always found to be effective in increasing innovative productivity.

The chapter then addresses a second issue, so far often neglected in the literature on spillovers: the positive externality generated by international technology flows will crucially depend on the destination country's ability to understand and exploit external knowledge. Such ability is a function of the country's past experience in research, an idea analogous to the concept of *absorptive capacity* introduced by Cohen and Levinthal (1990) in the context of firms' learning and innovation. The importance of incorporating this component into the analysis is suggested by the empirical dynamics of technological specialization emerging from the analysis of chapter 1.

The role of prior R&D experience in improving the ability of firms to understand and employ external knowledge has only been investigated in a few studies so far (see Griffith, Redding and Van Reenen, 2001; Griffith, Harrison and Van Reenen, 2003). The novelty here lies in the use of self-citations to measure the effect of absorptive capacity in enhancing the ability to benefit from spillovers. A self citation indicates that the firm did some research in the past and that it has now generated a new idea building upon previous research in the same or in a related technology field. As such, self citations are a clear indication of accumulation of knowledge internal to the firm.

The empirical results show that absorptive capacity increases the elasticity of a country's innovation to both national and international spillovers. However, its effect is different depending on the position of the country with respect to the world technological frontier: the larger the gap of a country with the technological leaders, the lower is its ability to absorb and exploit external knowledge, but the larger appears its potential to increase this ability.

CHAPTER 1

TECHNOLOGICAL SPECIALIZATION IN INDUSTRIAL COUNTRIES: PATTERNS AND DYNAMICS

1.1 Introduction

In the past two decades a series of studies on innovation has worked under the hypothesis that technological change is, to a large extent, a *cumulative* activity. In a well known paper, Dosi (1988) suggests the existence of technological paradigms. A '*technological paradigm*' defines contextually the needs that are meant to be fulfilled, the scientific principles utilised for the task, the material technology to be used. In other words, a technological paradigm can be defined as a pattern of solution of selected problems, based on highly selected principles derived from prior knowledge and experience. A technological *trajectory* can then be defined as technological progress along the economic and technological trade-offs defined by the paradigm.

This concept is very different from the identification of technology with information, easy to re-produce and re-use (Arrow, 1962). It recognises the highly differentiated nature of firms and of their search process (*firm-specific* nature of technology), which is pursued in areas that enable them to use and build upon their existing technological base (*local learning*). "*What the firm can hope to do technologically in the future is heavily constrained by what it was capable of doing in the past*"; put it simply, "(*technological*) *development over time ceases to be random, but is likely to be constrained to zones that are closely related technologically to existing activities*" (Dosi et al., 1990, pp. 84-85).

Another fundamental property of technological change relates to the forms and degrees of private *appropriability* of technological advances. As suggested by the classical and Schumpeterian traditions, varying degrees of private appropriation of the benefits of the innovation are both the incentive to and the outcome of the innovative process. In other words, each technology embodies a specific balance between its public good aspects and its private (i.e. economically appropriable) features. (Dosi et al., 1990, p. 89).

Finally, according to this view, differences in national rates of technological accumulation are influenced by strictly country-specific factors, among which: differences in the technological and institutional context, and technological externalities which act as a collective asset to single industries or group of industries within each national economy. "*These technological externalities affect the dynamics of innovation and imitation in each individual sector, and the overall pattern of technological accumulation in each country. (...) (They) are highly specific to particular spatial and institutional contexts; that they reproduce or even increase over time*" (Dosi et al., 1990, p. 107).

This last idea is close to those included in Arthur's work on *increasing returns, path-dependence* and *self-reinforcing mechanisms* in economics. It is well known that allocation problems with increasing returns tend to exhibit multiple equilibria. Arthur points out two new properties: *inflexibility* in that once an outcome (a dominant technology) begins to emerge it becomes progressively more "locked in"; and *non-ergodicity* in that historical "small events" are not averaged away and "forgotten" by the dynamics: they may decide the outcome.

The role of technological change and the possibility of increasing returns have become a key issue also in growth and trade theory. Both of them have become the focus of attention in recent empirical analyses of changes in the overall degree of countries' specialization and of the extent to which initial patterns of specialization persist over time. The theoretical literature on trade and growth typically yields ambiguous conclusions concerning both these issues. In particular, within the literature emphasising the endogeneity of technological change sector-specific learning-by-doing or localised knowledge flows are typically forces for persistence, while technology transfer across countries give rise to mobility (see, for example, Grossman and Helpman, 1991).

In endogenous growth theory models the creation of knowledge through private R&D yields positive external effects: part of the new knowledge adds to a public stock, accessible to all firms doing R&D themselves, thus reducing every firm's costs of future R&D. Over time, the public stock of knowledge grows, allowing more differentiated or higher quality products to be introduced without a continual increase in the amount of resources spent in R&D activities. This is referred to as knowledge spillovers, so called because the benefit of innovation accrues not only to the innovator, but "spills over" to other firms by raising the level of knowledge upon which new innovations can be based. Thus, knowledge spillovers serve as endogenous engine of economic growth.

The distribution of countries' output per capita and their comparative advantages are then determined by the process of technical progress in one country being independent from that in the others. Perfect technology diffusion (i.e. new ideas flowing as quickly to other countries as they flow within countries) favours the convergence of per capita output levels and leaves factor endowments as the sole determinants of trade patterns. However, if there are impediments to technology diffusion across national borders and the rate of knowledge spillover is much stronger within nations than across them[1], differences in levels of per capita output across countries will be persistent and the patterns of trade can exhibit path-dependence and lock-in (as a consequence of reinforcing effects characterising a country's technological change). Small initial inter-country differences lead then to divergence in specialization patterns and growth[2].

According to the theory the origin of persistence in the trade patterns of industrial countries may lie in the nature of technological progress or in the relatively stable position of advanced countries in the international economy. In the first case, stability in trade patterns arises as a consequence of technological progress being path-dependent and subject to localised knowledge spillovers, whereas in the second case knowledge spillovers may be pervasive and persistence is generated by stability in relative factor endowments. Understanding which of the two explanations applies is of primary importance, not least because they have different theoretical and normative implications.

If the actual specialization profile of a country is determined by its past strengths and weaknesses, industrial and technology policies targeted at selected industries and technologies in order to change the sectoral distribution of the country's comparative advantages would have lasting effects. Under the assumption that the government can identify the more promising technological trajectories, it can then pursue the deepening of specialization along those

[1] Jaffe (1989), Jaffe et al. (1993), Bottazzi and Peri (2003) suggest this, among others. Keller (2000) finds evidence pointing to the relevance of geographical proximity, but also shows that the detrimental effect of geographical distance on international technology diffusion has fallen by about 20 percent over the period 1970-1995.

[2] If a country acquires a temporary advantage in an R&D intensive sector, it can innovate in that sector at a faster rate than other countries. This is because the knowledge base on which domestic firms build their innovations grows faster than anywhere else, given that it cannot quickly spread to foreign competitors. Hence the country can build on an initial advantage, eventually developing a position of enduring comparative advantage.

technologies, by shifting resources towards them.

Therefore, whether international trade patterns persist or exhibit mobility over time (and whether there is increasing or decreasing specialization over time) is an empirical question. Empirically, if the explanation of the stability in trade specialization were based on the nature of technology, then persistence should be particularly pronounced in technological specialization, where the positive external effects in the form of knowledge spillovers from R&D have their most direct and strongest impact.

This chapter provides an empirical assessment of the relevance of the hypothesis of path-dependence in the sectoral technological development of open economies, a hypothesis related to the supposedly localised nature of external economies from knowledge creation in the innovation literature and in recent models of endogenous technological change, specialization and growth.

In the analysis I represent a country's pattern of international specialization in technologies through a distribution of relative innovative output shares across technologies and then refer to its dynamics as the evolution of the entire distribution over time. This very general specification is consistent with a wide range of possible technology dynamics and allows determining the degree of persistence versus mobility in patterns of international technological specialization from the observed data. It also allows determining whether the observed dynamics is consistent with path-dependence in its strongest version (i.e. reinforcing effects leading to polarization and, possibly, lock-in).

This purpose is here achieved applying the dynamic tools offered by distribution dynamics modelling (Quah, 1993a, 1996a, 1996b, 1996c) to the analysis of the evolution of the technological specialization profile of industrial countries in the last two decades. This approach is appropriate to study the evolution of a country's specialization pattern as a process where the state of the system determines the probability of the next action (see Arthur, 1994, pp. 119-120).

Countries do not show increasing specialization in a limited subset of technologies (a polarization towards the extreme values of the distribution representing the specialization pattern). However, for most countries the relative stability in the degree of specialization hides significant intra-distribution mobility. In particular, high specialization levels are not persistent in time; rather, they revert towards lower levels, the reversion being faster and more pronounced for smaller countries (i.e. countries with higher overall degree of specialization). By contrast, higher persistence is found in situations of initial complete or high de-specialization.

These results have implications on the relationship between technological change, growth, and trade of countries, and complement the studies on the empirical patterns of trade dynamics. They are not in line with a theory of technological accumulation and path-dependence, and are consistent with the findings in Stolpe (1995), which cast doubt on the often alleged causality from hysteresis in technology to hysteresis in trade specialization. The results might also help explaining the finding of high mobility in trade patterns emerging from recent empirical studies on trade dynamics[3].

The chapter is organized as follows. Section 2 provides a brief review of the literature on path dependence and technological specialization. Section 3 shows how technology dynamics can be seen as a recontracting process whose stationary distribution is characterised by strong specialization/de-specialization if national externalities are strong. Section 4 defines the technological specialization profile of a country and explains how it can be measured. Section 5 studies the evolution of the overall degree of technological specialization through changes in an inequality index and the non-parametric estimation of the density functions representing the specialization patterns of ten OECD countries. The same section then studies intra-distribution

[3] See Proudman and Redding (1998, 2000) and Brasili et al. (2000).

movements (i.e. changes in time of a country's specialization in a particular technology field) through the estimation of Markov stochastic kernels and transition probability matrices. Section 6 concludes. Empirical methods are presented in the Appendix.

1.2 Increasing returns, path-dependence and technological accumulation

During the past twenty years economists have exhibited a growing interest in the idea of increasing returns on a wide range of subjects, including the spatial location of production, the development of international trade, the causes of economic growth and the emergence of new technologies[4].

Arguments about technology have provided the most fertile ground for exploring the conditions conducive to increasing returns. As Brian Arthur (1994) and Paul David (1986) have stressed, under conditions often present in complex, knowledge-intensive sectors, a particular technology may achieve a decisive advantage over competitors, although it is not necessarily the most efficient alternative in the long run. Once an initial advantage is gained, positive feedback effects may lock in this technology, excluding competitors.

This last idea is illustrated in Arthur's model of competing technologies (Arthur, 1989). The author explores the dynamics of allocation under increasing returns, within a model where agents choose between technologies competing for adoption and where each technology improves as it gains in adoption. It shows that the economy, over time, can become locked-in by "random" historical events to a technological path that is not necessarily efficient, not possible to predict from usual knowledge of supply and demand

[4] It should be noted that the ideas developed in this research are not entirely new. The concept of increasing returns received attention already in the work of Alfred Marshall.

functions, and not easy to change by standard tax or subsidy policies. Rational expectations about future agents' technology choices can exacerbate this lock-in tendency.

In the presence of positive feedback, the probability of further steps along the same path increases with each move down that path. This is because the relative benefits of the current activity compared to once-possible options increase over time or, alternatively, the costs of switching to some previously plausible alternative rise. As a consequence, actors have strong incentives to focus on a single alternative and to continue down a specific path once initial steps are taken in that direction. The general point is that sequencing is critical in such processes. Earlier events matter much more than later ones, hence different sequences may produce different outcomes: history matters.

The same ideas underlie the theory of technological accumulation, according to which technological change exhibits reinforcing effects due to geographically bonded and sector specific learning by doing or knowledge externalities. According to this view a country's ability to innovate and its pace of technological progress in a field depend on its historical leads and lags. As a consequence, "*international patterns of technological advantage, having been established, will remain relatively stable over periods of ten or even twenty years, under the assumption that only the emergence of new technological paradigms and industries can, in the long term, generate important changes in the specialization trajectories of both firms and countries*" (Cantwell 1989).

Cantwell (1989), Archibugi and Pianta (1992a, 1992b, and 1994), Stolpe (1995) and recent studies by Amendola et al. (1998) and Laursen (2000) have analysed the technological specialization profiles of advanced countries, trying to establish whether there is evidence of their stability over time. With the exception of Stolpe (1995) they all employ static methods of analysis[5]. The

[5] Cantwell (1989), Amendola, Guerrieri and Padoan (1998) and Laursen (2000) employ the

findings these studies reach are mixed, depending on the specific sample and measure of specialization used. However, an overall picture of stability in technological specialization patterns emerges, with the exception of France and the UK and with the US showing a tendency towards an overall increase in its degree of specialization. This result is somewhat mitigated in Stolpe (1995) and Laursen (2000), who provide a comparison between stability in technology and trade specialization patterns and find the first to be lower than the latter, thus suggesting relatively low persistence in technological specialization.

1.3 Technology dynamics and recontracting processes

The evolution of technological specialization and the possibility of geographically bounded self-reinforcing effects in innovative activities can be illustrated with reference to a *recontracting process*, which considers a total allocation of fixed size divided among K categories. Transitions of units between categories are possible, with probabilities that depend, in general, on the market shares or numbers in each category. Thus self-reinforcement is possible.

The simplest way to illustrate some of the basic implications of national external effects for the dynamics of technological specialization is within a stylised model where labour is the only factor employed in the production of new knowledge, the labour market is perfectly competitive, and the economy's R&D labour force is given and equal to $2N$.

Galtonian regression model, which may suffer of the well-known Galton's fallacy (Hotelling 1932; Friedman 1992; and Quah 1993b). In Archibugi and Pianta (1992a, 1992b, and 1994) changes over time of the profile and degree of specialization are analyzed by looking, respectively, at the correlation coefficients of specialization vectors at different time periods and at the evolution of the chi-square index over time. Only Stolpe (1995) uses the same methodology employed here to assess the evolution of patterns of technological specialization in OECD countries, albeit with an emphasis on the distinct dynamics within individual industries.

To keep things simple, consider the case where K, the number of technological fields in which the workers can be employed, is equal to 2 (field 1 and field 2). Workers decide with which technology to work and can move from one to the other. In this setting, a comparative advantage in new knowledge production is due to higher labour productivity in one of the two technology fields and increasing returns to scale can take the form of a positive externality for all innovators active in a field.

Under the assumption that transitions can be made only one unit at a time, the resulting reallocation of labour between the two technologies can be characterised as a stochastic recontracting process similar to the one studied by Weidlich and Haag (1983), and later taken up by Arthur (1994). This can be modelled as a one-dimensional stochastic process in time, because the relative shares of the two technological fields in the total innovative activity of the model economy require one state variable only. Given the configuration $\{n_1, n_2\}$ at a given time t, consisting of the numbers n_1 and n_2 of workers employed with technology 1 and 2, respectively, the state variable n is such that:

$$\begin{array}{lll} n_1 + n_2 = 2N & \text{and} & n_1 - n_2 = 2n \\ n_1 = N + n & \text{and} & n_2 = N - n \\ -N \le n \le N & \text{and} & 0 \le n_1, n_2 \le 2N \end{array} \qquad (1)$$

Let $p_{12}(n)$ denote the probability that a worker moves from technology 1 to technology 2, and $p_{21}(n)$ the probability that a worker moves from technology 2 to technology 1, in unit time (i.e. $p_{12}(n)$ and $p_{21}(n)$ are the individual transition probabilities) and consider a population of economies (each consisting of a labour force equal to $2N$). The function

$$p[n_1, n_2; t] \equiv p(n;t) \qquad (2)$$

denotes the probability that one sample economy has the configuration $\{n_1, n_2\}$ at time t. Of course, the condition

$$\sum_{-N}^{N} p(n;t) = 1 \qquad (3)$$

has to be satisfied at all times.

The equation of motion for $p(n;t)$ can be derived considering that the individual transition probabilities induce "nearest neighbour" transitions of the configuration $\{n_1,n_2\}$ only. The transition $\{n_1,n_2\} \rightarrow \{n_1+1,n_2-1\}$, or equivalently $n \rightarrow n+1$, is effected by a transition from technology 2 to technology 1 by one of the n_2 members working with technology 2. Analogously, a transition $\{n_1, n_2\} \rightarrow \{n_1-1, n_2+1\}$, or equivalently $n \rightarrow n-1$, is effected by a transition from technology 1 to technology 2 by one of the n_1 members working with technology 1.

Correspondingly, the transition probabilities for the whole configuration are given by[6]

$$\begin{aligned}
&w(n \rightarrow n+1) \equiv w_{\uparrow}(n) = n_2 p_{21}(n) = (N-n) p_{21}(n) \\
&w(n \rightarrow n-1) \equiv w_{\downarrow}(n) = n_1 p_{12}(n) = (N+n) p_{12}(n) \\
&w(n \rightarrow n') = 0 \quad \text{for} \quad n' \neq n \pm 1
\end{aligned} \qquad (4)$$

and the equation of motion for $p(n;t)$ is equal to

[6] The model can be generalised to multiple transitions for the case in which m members of the labour force simultaneously change the technology they work with. Methods of solution become more complicated and are discussed in Weidlich and Haag (1983).

$$\frac{dp(n;t)}{dt} = [w_{\downarrow}(n+1)p(n+1;t) - w_{\downarrow}(n)p(n;t)] + [w_{\uparrow}(n-1)p(n-1;t) - w_{\uparrow}(n)p(n;t)] \tag{5}$$

The interpretation of equation (5) is the following: the change in time of the probability of state n is due to the probability fluxes into and out of state n, that is the probability fluxes from or towards states (n+1) and (n–1). Equation (5) has also an important property: an arbitrary initial distribution finally develops into a stationary equilibrium distribution $p(n)$, that is the recontracting process shows convergence in distribution (Weidlich and Haag, 1983: p. 9).

In an interesting illustrative example, Weidlich and Haag (1983) show that the stationary distribution can show either a centralising tendency or a bimodal shape, where the modes are towards the end states of complete specialization into one of the two technologies, depending on the strength of the national, technology specific external effects. In this example, the probabilities of a worker moving from one technology to the other are:

$$\begin{aligned} p_{21}(n) &= \nu \exp(\delta + \tilde{\kappa} n) = \nu \exp(\delta + \kappa x) \\ p_{12}(n) &= \nu \exp[-(\delta + \tilde{\kappa} n)] = \nu \exp[-(\delta + \kappa x)] \end{aligned} \tag{6}$$

where $\kappa = N\tilde{\kappa}$, $x = n/N$ $(-1 \leq x \leq 1)$, and δ, κ and ν are parameters whose effect on the transition probabilities can be described as follows:

a) The parameter δ allows for a "preference bias". A positive δ increases the probability that a worker moves from technology 2 to technology 1, and reduces the probability of moving from 1 to 2. The opposite happens for a negative δ. In the present setting the parameter δ could signal that workers are more productive in one of the two technologies, giving the country an ex-ante comparative advantage in that technology[7].

b) The parameter κ is the national external effect[8]. A positive κ enlarges the transition probability in favour of the technology with the largest share of workers and reduces the transition probability in the opposite direction. This effect grows for growing $|x|$ or for growing imbalance in the shares of the two technologies. Through the term κx, the probabilities $p_{ij}(n)$ depend on the present configuration.

c) The parameter ν determines the frequency of switches, or the time scale in which changes of technology by workers occur.

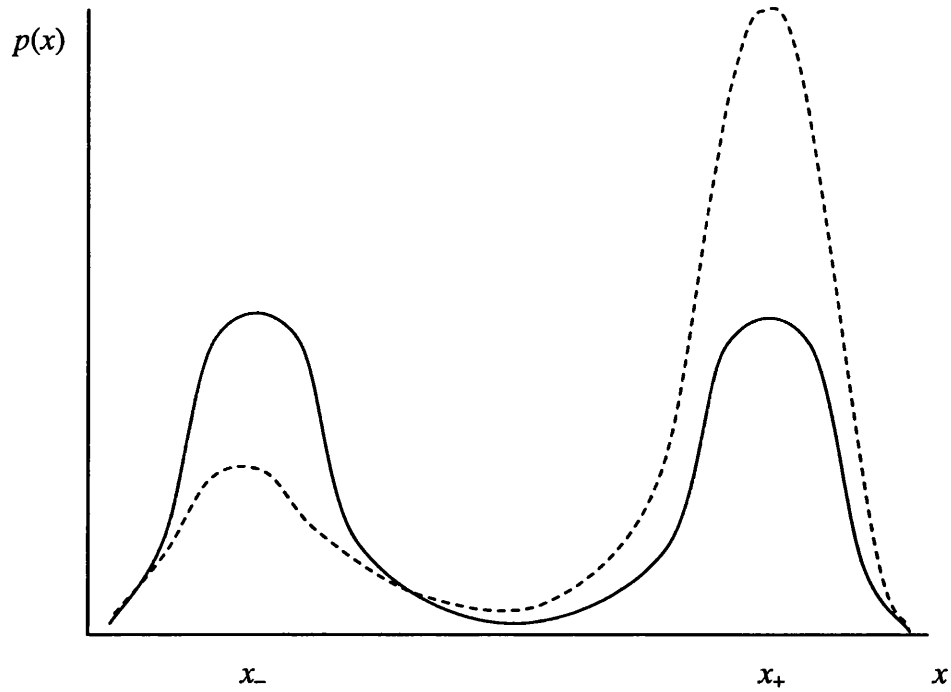
From equation (6), transition probabilities for the whole configuration can be derived as in (4), which substituted into the equation of motion for $p(x;t)$, allow obtaining the stationary distribution $p(x)$ (see Weidlich and Haag, 1983: section 2.4). Its properties depend on the chosen values of κ and δ, and are independent of ν. In the absence of conformity, a larger population in one technology increases the chance of switches to the other; hence there is a centralising tendency. This is offset by the conformity effect, which reinforces a concentration of one type.

When κ is small ($\kappa < 1$), centralisation dominates and the stationary distribution is unimodal. The independent preferences of the individual workers, described by δ, play the main role. For $\delta = 0$, the most probable outcome (the mode of the distribution) is one with $n = 0$, or $n_1 = n_2$, i.e. equal specialization in technologies 1 and 2. When $\delta > 0$ (the case of comparative advantage in technology 1), the mean value of the distribution is shifted to the right ($E(n) > 0$), but the shape of the distribution remains approximately the same.

[7] Indeed, higher productivity in, say, technological field 1 would imply higher wages from a job in technological field 1, hence higher probability of switching from field 2 to field 1.
[8] Arthur (1994) refers to this as the "conformity" effect, while Weidlich and Haag (1983) call it "adaptation" parameter.

As κ increases ($\kappa > 1$) the distribution bifurcates and becomes bimodal, with maxima corresponding to the relative prevalence of one technology over the other. For $\delta = 0$, the two modes have equal height: equal specialization in the two technologies is unlikely, while specialization in either of the two is most probable. In the case when $\delta > 0$ a large probability peak corresponding to a value x_+ close to the boundary $x = +1$ is found, while there is still a now small probability peak at a value x_- close to the boundary $x = -1$ (Figure 1.1). This second peak implies that, because of previous history, a stable configuration x_- exists in contradiction to the sign and magnitude of the individual preference parameter δ of all members of the economy.

Figure 1.1 Stationary distributions, $p(x)$, with high localised externality effect (κ>1), when δ=0 (—) and δ>0 (- - -)



To sum up, the recontracting process presented in this section formalises the idea that in the presence of strong national externalities (when international spillovers are absent or weak) countries should display the tendency to specialize in selected technologies, either because they have a comparative advantage in such technologies or because of hysteresis. In a setting with more than two technologies one should then expect the specialization pattern of countries to evolve towards areas of clear specialization and areas of clear de-specialization, the tendency being stronger for countries with already unbalanced initial specialization patterns. This means one should observe increasing overall specialization in technologies.

Finally, and related to the considerations above, established technological specialization profiles of countries should display persistence, with particularly low mobility between high specialization and low de-specialization states (i.e. between the peaks of the stationary distribution).

1.4 Empirical modelling of technological specialization

In order to measure a country's international (horizontal) specialization in technologies (i.e. in the production of new knowledge across technologies) I need disaggregated and international comparable data on knowledge output. For this purpose, I use patent data from the EPO-CESPRI database, which covers all patent applications filed at the European Patent Office (EPO) between 1978 and 1996. Patents are, indeed, the only available indirect evidence of technological activity offering a detailed breakdown by technological areas for a large number of countries and for long time series. Hence, they will be used here to characterise countries' distribution of research output across technologies: this will summarize the technological frontier of a country and its pattern of technological specialization at a specific point in time.

Specialization of a country in a technology relative to other countries is captured by world shares rather than absolute levels. Normalizing the shares by the average share across technologies ensures the comparability of various technologies and countries. The resulting measure is similar to the one used by Proudman and Redding (1998). Horizontal specialization is then given by the dispersion of output shares across technologies. Increasing horizontal specialization translates into a more unequal distribution of output shares within a country and across technologies. Likewise, increasing horizontal diversification translates into a more equal distribution of output shares.

1.4.1 Patent data

A patent is a legal title granting its holder the exclusive right to make use of an invention for a limited period and in a limited area[9], by stopping others from making, using or selling it without authorisation. The stated purpose of the patent system is to encourage technical progress by providing a temporary monopoly for the inventor and also by forcing the early disclosure of the information necessary for the production (if a product) or the operation (if a process) of the object of the patent. On one side, the exclusive right to commercially exploit an invention represents an incentive for firms to engage into (and finance) research and development; on the other side, patents play a major role in the transfer and dissemination of new technologies.

To be granted, a patent has to satisfy three criteria: *novelty*, *inventive step*[10] and *industrial applicability*. A substantive examination is carried forward to

[9] A European patent is valid for 20 years and is awarded protection in the designated contracting states. Further details are given in footnote 13.

[10] An invention is considered new if it is not part of the state of the art; and it is considered as involving an inventive step if it is not obvious to a person having ordinary skill in the area of technology related to the invention.

verify that those criteria are met. This may require quite a long time[11], so that when patents are employed in economic analysis, they are usually classified according to their application or priority date[12], on the grounds that it is closer to the time the innovation was accomplished. The procedures and criteria for granting patents may greatly differ across countries. It follows that cross-country comparisons are most reliable when using international patenting or patenting in one country. The US patenting statistics have most often been used in applied economic studies on patent data as the privileged available source of information for international comparisons on technological activities and specialization patterns. This is certainly because of the rigorous and fair screening procedures applied by the US Patent Office, but also because of the strong incentives for firms to get patent protection for world class technology in the world's largest market. As the EPO was established in 1977 another potential source became available. Now the EPO patent statistics include a sufficiently long time series of observations. These can be used on the grounds that the EPO too adopts unified granting criteria and procedures and that Europe represents a large market as a whole, where firms compete across their national borders and which attracts further competition from firms established in other markets.

Using patenting at the EPO to explore a country specialization profile has an additional advantage: all the firms patenting at the EPO are patenting abroad[13],

[11] On average, it takes 44 months to obtain a European patent.

[12] In each patent office in which an application is filed, an application number and an application date are assigned to the document. The first filing application is considered the priority application and the date of this application is the priority application date or, simply, the priority date. Usually firms apply for a patent at the EPO after they already applied for patent protection at their national patent office: the priority date in EPO patent documents refers to this first filing.

[13] The EPO was founded on the basis of an agreement among 13 European countries: Austria, Belgium, France, West Germany, Greece, Italy, Liechtenstein, Luxembourg, the Netherlands, Spain, Sweden and the United Kingdom. Six more states have become members later (Cyprus, Denmark, Finland, Ireland, Monaco and Portugal). A single application at the EPO can potentially be extended to all the member countries, and on average, the number of contracting

so that the analysis based on this data is not greatly affected by the "domestic market bias". Domestic patenting is an unreliable indicator of a country's specialization, as it is distorted by a large number of innovations of low importance, which are not extended abroad. This is partially due to the presence of individual inventors. Each patent institution receives an above average number of domestic applications in the selected technological fields where individual inventors are usually more active (e.g. consumer goods), but only occasionally do they extend their patents abroad[14]. Another potentially relevant effect is related to the tendency national firms have to patent their inventions (and hence protect their domestic market share) also in fields where they have not developed a world class technological capability, as an additional defence against competitors. On the contrary, they will be willing to patent abroad inventions of greater quality, for which there is a higher probability of economic returns that will justify the additional protection expenses.

Such characteristics result in a much less clear specialization pattern, so that the areas where a country is internationally strong can be hardly identified within the vast and more uniform domestic patenting activity. This is likely to have adversely affected the characterisation of the technological specialization profile of the United States emerging from previous studies, which was mostly based on US domestic patenting. The resulting specialization profile does not appear as an accurate description of the areas of technological strength and weakness of the US in the international markets[15]. The analysis of data on

states designated for protection is about 8 per patent.

[14] Although individual inventors show a lower tendency to patent abroad, they still do. Nevertheless, this will not affect the analysis that follows, since individual inventors have been excluded from the database.

[15] It has been observed elsewhere that the specific strengths and weaknesses of the US differ substantially in domestic and in external patenting, with no correlation emerging between data from the US patent office on the one hand and from the EPO, France and Germany patent offices, on the other. The specialization profiles emerging from different foreign markets are consistent, however, with correlation coefficients always higher than 0.7 (Archibugi and Pianta, 1992a and 1992b).

patent applications from US firms at the EPO can therefore prove particularly useful.

Ideally, we would like patent statistics to provide a measure of the output of the innovative activities, those activities that lead to a reduction in the cost of producing existing products or to the development of new products and services. This would provide us with a direct reading of the rate at which the potential production possibility frontier is shifting outward. Indeed, it has been widely recognised that patent statistics offer a potentially very rich source of empirical evidence on questions related to the structure of technology and how it changes over time across countries, industries and firms[16]. However, it is also often reminded that there are some major problems in using patents for economic analysis: *intrinsic variability* and *classification*.

The first problem refers to the stochastic fluctuations in the propensity to patent and the variation in importance of individual patents. Even fields to which similar amounts of innovative resources are devoted can show great differences in their level of patenting[17] and firms themselves may have different propensities to patent[18]. As a consequence, not all innovations are patented, but even those that are patented differ greatly in their "quality" (Schankerman and Pakes, 1986). Because patent rights are seldom marketed, there are only few sources of information on their economic value, among

[16] See, for example, Pavitt (1988), Griliches (1990).

[17] In the field of chemicals, for instance, patent protection is systematically used to appropriate returns from innovative activities; patents, instead, do not measure satisfactorily advances in software technology, where the practice of protecting innovation through patents is of recent origin.

[18] Scherer (1983), using US patent data, found that most of the variation between firms in the propensity to patent was to be explained by the extent of their research effort (as measured by R&D expenditure). It nevertheless seems reasonable to assume that, even allowing for inter-firm intra-field differences in the propensity to patent of the firms of a given country, their variance is systematically lower than that for inter-field differences. This will justify the aggregation of firm-level data into industry-level data for each country and the derivation of the specialization index based on the latter. Since this index is examined separately for each country, no assumption is required about international differences in the propensity to patent in a given technological field.

these the results of direct surveys of patent owners, data on renewals[19] and citations, data on some value-denominated variables, such as the profits or stock market value of the patenting firm. Unfortunately we do not have yet a universally acknowledged good procedure for weighting patents appropriately so to account for their intrinsic value. Nevertheless, also R&D projects (the alternative and widely used indicator of innovation activity) have the same drawback (Freeman, 1982).

The problem of classification is primarily technical and refers to the difficulties in allocating patent data, organised by firms or patent classes, into economically relevant industries or product groupings. Each Patent Office classifies patents into many classes and sub-classes, so to ease the search for prior art[20]. The resulting classification system is based primarily on technological and functional principles and cannot easily be related to product or industry classifications[21]. I refer to Griliches (1990) for an extended discussion on this problem. For what matters here, it should be noted that, the analysis being on technological specialization, this drawback in the use of patent statistics is not so relevant in what follows. Still, it is worth keeping in mind that the classification of patents must be distinguished from the industry where the firm that has applied for the patent is active. Hence, in the following pages I shall use the term "field" or "technology" to remind I am not referring to production sectors or product groups.

1.4.2 Specialization in technologies: definition and measurement

To characterise the extent of specialization in a technology, previous

---

[19] Schankerman and Pakes employ data on renewals in a series of papers. They estimate models that allow them to recover the distribution of returns from holding patents at each age over their lifespan. See , for example, Schankerman and Pakes (1986).

[20] A brief description of the structure of the International Patent Classification is reported in Appendix A: The International Patent Classification.

[21] For example, a subclass relating to the dispensing of solids contains patents both on manure spreaders and toothpaste tubes. See Schmookler (1966) for other examples of this kind.

studies[22] have employed the so called Technological Revealed Comparative Advantage (*TRCA*) index, computed in the same way as Balassa's (1965) index of Revealed Comparative Advantage (*RCA*) used in trade theory. The *TRCA* index is defined as a country's share of EPO patenting in a technology, relative to its share of EPO patenting in all technologies,

$$TRCA_{ij} = \frac{P_{ij} / \sum_i P_{ij}}{\sum_j P_{ij} / \sum_{ij} P_{ij}} \tag{7}$$

where P_{ij} denotes the number of EPO patents of country i in technology j.

The *TRCA* yields information about the pattern of international technological specialization insofar as it evaluates a country's patenting share in an individual technology, relative to some benchmark: the country's share of total patenting. A value of $TRCA_{ij}$ above unity indicates that country i is comparatively advantaged or specialized in technology j.

Although the *TRCA* index varies around unity, it suffers from the disadvantage that its arithmetic mean across technologies is not necessarily equal to one. The numerator in equation (7) is unweighted by the proportion of total patenting accounted for by a given technology, while the denominator is a weighted sum of patenting shares in all technologies. As a consequence, difficulties can arise, especially when constructing a *TRCA* index for small countries, which typically apply only for low numbers of patents at the EPO.

Small and open economies are usually more specialized and internationalised than large economies. Indeed, relatively small countries are to some extent forced to specialize in selected niches, because they lack the resources and technological expertise needed to carry out expensive R&D that entails risks and gives uncertain outcomes. In such a setting, the pattern of

[22] See, for example, Soete (1981) and Patel and Pavitt (1991).

patenting of a small open economy will often be characterised by high patenting shares in few technologies, each of which accounts for a small share of total patenting. This implies that small economies will be often characterised by mean value of *TRCA* above one, and that their technological advantages or disadvantages, as measured by the *TRCA* index, will be characterised by substantial inter-technology variation. Some small countries may appear among the highest centres in the hierarchy for a particular technology, even though they are not among the most important centres. Hence, some very low or high values of the *TRCA* index may be misleading for the purpose of cross-country comparisons in a technology. Furthermore, mean values of *TRCA* may change over time, so that an economy exhibits changes in its average extent of specialization over time. This is yet another drawback of the *TRCA* index, since in analysing the change in a country's patterns of international technological specialization across technologies it is desirable to abstract from variations in its average extent of specialization[23].

Proudman and Redding (1998 and 2000), have recently adopted a modified version of that *TRCA* index, according to which an economy's share of patenting in a given technology is evaluated relative to a different benchmark, which is its *average* patenting share in all technologies:

$$RTA_{ij} = \frac{P_{ij} / \sum_i P_{ij}}{\frac{1}{N}\sum_j (P_{ij} / \sum_i P_{ij})} \qquad (8)$$

where N is the total number of technologies.

By construction, for each country the mean value of *RTA* across

[23] Various studies have used a normalised version of the *TRCA* index, which may be easier to interpret, but which nevertheless suffers from this same problem. This is calculated as: $BNRTA_{ij} = (TRCA_{ij} - 1)/(TRCA_{ij} + 1)$. Hence $-1 \leq BNRTA_{ij} \leq 1$ and positive and negative values indicate areas of country's specialization and comparative disadvantage, respectively.

technologies is constant and equal to one[24]. Again, a value of RTA_{ij} above unity indicates that country *i*'s share of patenting in technology *j* exceeds its average share in all technologies: that is, country *i* specializes in technology *j*. Note that *RTA* ranges between 0 and N[25], hence it is asymmetric as the *TRCA* index, with which it is perfectly correlated.

1.5 The evolution of the technological specialization profile

The pattern of international technological specialization at any one point in time can be characterised by the distribution of *RTA* across technologies. Hence, evaluating the dynamics of patterns of international technological specialization over time requires an analysis of the evolution of the entire cross-section distribution of *RTA*. This involves two different, but related issues. On one hand, there is the issue of the changes in the overall degree of international technological specialization, which may be evaluated by analysing the evolution of the *external shape* of the *RTA* distribution. Do we observe an increasing specialization in a limited subset of technologies (a polarization of the *RTA* distribution towards extreme values), or has the degree of specialization remained broadly unchanged?

On the other hand, there is the issue of persistence versus mobility in international technological activities. This addresses questions related to the

[24] It can be easily shown that:

$$RTA_{ij} = TRCA_{ij} / \frac{1}{N} \sum_j TRCA_{ij}$$

The TRCA measure is normalized by its cross-sectional mean in order to abstract from the changes in the average extent of specialization that it is subject to. In this way, it is always possible to follow movements of a country's specialization in a field with respect to its average specialization level.

[25] Note that the maximum value *TRCAij* can take varies both in time and across countries. *TRCAij* is highest when country *i* only applies for patents in sector *j* and no other country does. In this particular case, *TRCAij* reduces to $(\sum_{ij} P_{ij}) / P_{ij}$, which differs across countries, while *RTAij* is equal to N.

intra-distribution dynamics, such as: what is the probability that a technology moves from one quartile of the *RTA* distribution to another? Are all technologies with RTA_{ij} > (<) 1 at time *t*, still specialized (de-specialized) at time *t+k* (k ≥ 1)? If not, at what level of specialization is the greatest degree of mobility observed, and how far are those technologies moving towards low (high) values of *RTA*?

The present study mostly differs from previous empirical studies on technological specialization in that it directly addresses these two fundamental issues[26]. Indeed, Arthur has emphasised that a dynamic approach is needed to allow for the sequence in which actions occur and to see if these are affected by the numbers or proportions of each alternative present at the time the action is taken (Arthur, 1994: p. 119). For these reasons, the evolution of the *RTA* distribution over time is here modelled adopting a distribution dynamics approach, a technique recently developed by Quah (1993a, 1996a, 1996b, 1996c) to analyze income convergence in the cross-country growth literature. This empirical methodology is more informative than the ones adopted in

[26] Many of these studies employ the *Galtonian regression model* (see footnote 5): for each country the value of the specialization index at time t_1 is regressed on a constant and the value of the same index at the earlier time t_0 in a simple cross-technology regression. The sign and magnitude of the estimated slope coefficient give information on the type of correlation between past and actual specialization profiles. An estimated slope coefficient equal to one implies that technologies retain the same proportional position (i.e. they remain advantaged or disadvantaged exactly as they were in the previous period), with consequently no change in their ranking. An estimated slope coefficient greater than one represents the case of a proportional shift in which already advantaged (disadvantaged) technologies tend to become even more advantaged (disadvantaged). Finally, when the estimated slope coefficient is smaller than one, disadvantaged technologies improve their position, while advantaged ones slip back: this phenomenon is known as *regression towards the mean*. There is an important reason to be cautious when using this approach, which relates to the well-known *Galton's fallacy* (Hotelling, 1932; Friedman, 1992; Quah, 1993b). Regression towards the mean cannot necessarily be interpreted as a tendency of the extremes of the cross-technology distribution of the technological specialization index to converge towards the centre (Hotelling, 1932). While observations at the margins often go towards the centre, those in the centre may also move towards the margins, some going up and others going down. Since the positive and negative deviations cancel in averaging, while for observations at the extremes the only possible motion is towards the centre, we observe a reduction in the spread of the observations, which nevertheless does not imply at all they are indeed converging towards a common centre. It only proves that the observations in question have a tendency to wander about.

previous empirical studies because it exploits both the cross-sectional and the time series variability in the data[27].

The analysis presented in the following pages will address the first issue above (*changes in the external shape of the distribution*) by estimating the distribution of *RTA* across technologies for each country. It will then shed light on the *persistence vs. mobility* issue, where persistence will be interpreted as a measure of the probability of remaining in the state in which a country initially is. Namely, if a country is specialized in a technology, the question is: what is the probability that it remains specialized as time goes by. All the probability density functions, Markov stochastic kernels and transition probability matrices, presented in the following sections, have been estimated from EPO data on 118 3-digit technologies, classified according to the International Patent Classification (IPC) system, using Danny Quah's econometric shell TSRF.

In what follows I shall restrict the analysis to the period 1982-1996, thus dropping the very first years of activity of the EPO, which were characterised by a relatively low number of applications. The countries analyzed here include the first ten countries for number of applications at the EPO: Austria, France, Germany, Italy, Japan, Netherlands, Sweden, Switzerland, United Kingdom, United States. The number of patent applications filed each year at the EPO varies widely across countries. On average, each year US firms apply for 13629 patents, a number well above the average number of applications coming from each of the other countries. The other countries with a relatively high level of patenting activity are: Germany (9094 applications each year, on average), Japan (8724), France (3600), the UK (2719), Switzerland (1684), Netherlands (1554), Italy (1467), Sweden (765) and Austria (397). These numbers imply a share of about 30 percent of all the applications at the EPO

[27] Stolpe (1995) is the only other study in the field that adopts the same methodology.

for the US throughout the whole period. Germany and Japan have a share fluctuating around 20%; France has a constant 8% share, while the UK has a share declining from 8% in 1982 to 4% in 1996. All the remaining countries have a share lower than 5 percent.

1.5.1 Changes in the degree of specialization

A measure of specialization that can be borrowed from the inequality literature is the widely used Gini coefficient. To derive it for a country's specialization in technologies, first rewrite the *TRCA* index in country i and technology j as the ratio of class j's patenting share in country i over technology j's share of total patenting[28]. Then construct the Lorenz curve as follows: rank the *TRCA* index in ascending order, then plot the cumulative of the numerator on the vertical axis against the cumulative of the denominator on the horizontal axis[29]. The Gini is equal to twice the area between the 45-degree line and the Lorenz curve. If the technological structure of country i matches the world technological structure, the Lorenz curve will coincide with the 45-degree line and the Gini coefficient will be zero. The higher the Gini the more specialized is a country.

The evolution of the Gini coefficient for the analyzed countries is reported in Figure 1.2. The degree of specialization of the G5 countries is relatively low (the Gini coefficient is always below 0.3 except, but only occasionally, for Japan) compared to that of the other countries, which is also characterised by wider fluctuations[30].

---

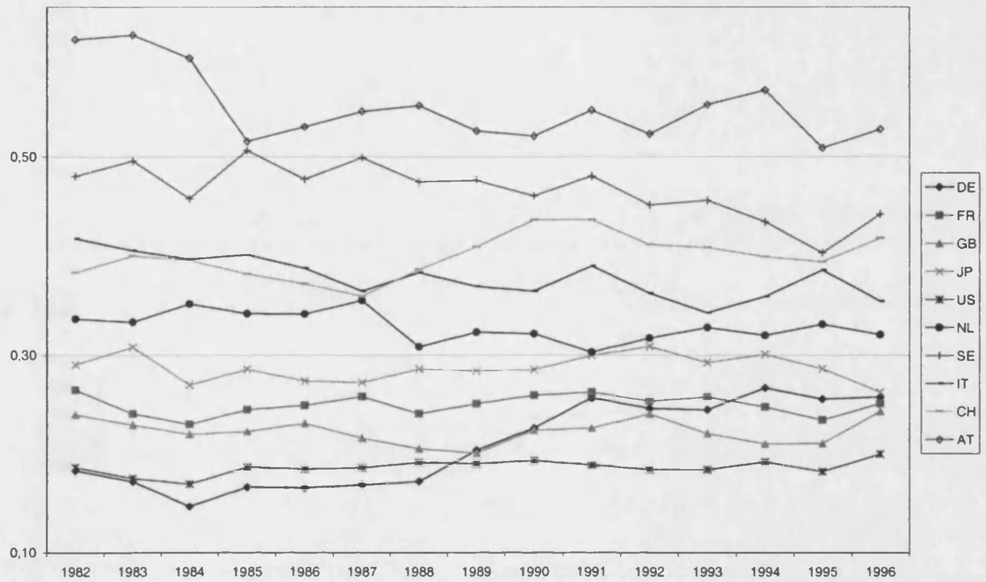
[28] Hence:

$$TRCA_{ij} = (P_{ij} / \sum_j P_{ij}) / (\sum_i P_{ij} / \sum_{ij} P_{ij})$$

[29] Note that by constructing the Lorenz curve in this way I am comparing the distribution of country i's patenting across technologies to the distribution of the total patenting across techologies and not to the uniform distribution, as it is usually the case (see Amiti, 1999).

[30] Note, however, the significant change in the Gini coefficient for Germany between 1989 and 1991. This is probably a consequence of the unification with East Germany.

Figure 1.2 The evolution of the Gini coefficient for country specialization



Note, that the analysis on patent applications at the EPO shows a tendency towards a declining degree of technological specialization in OECD countries, contrary to previous findings based on US patent data (e.g. Cantwell 1989). This tendency appears stronger for countries with higher initial degree of specialization. Indeed, a simple cross-section linear regression of the change in the Gini coefficient during the sample period on its initial value reveals a strong and significant negative effect (see regression (3) in Table 1.2). This in terms of shares from one technology to another, how does this affect the inequality measures obtained above? The Gini coefficient places a rather curious relative value on changes that may occur in different parts of the distribution. Hence, a transfer from a more specialized technology to a less specialized one has a much greater effect if the two technologies are near the middle rather than at either end of the spectrum[31]. The main problem relates to the difficulties in identifying and evaluating intra-distribution movements, which might also generate changes in the shape of the overall distribution.

To take care of this problem, recall that the pattern of technological specialization can be represented by the distribution of the *RTA* index across technologies, hence it can be estimated from the data for each country at different time periods. This has been done for all the countries by pooling the observations into three sub-periods: 1982-1986, 1987-1991, 1992-1996 (see Figure 1.3). All the densities have been estimated by Gaussian kernel smoothing, taking non-negativity into account and following the procedure and automatic bandwidth choice from Silverman (1986: 2.10 and 3.4.2).

The *United States*, *Germany*, *France* and the *UK*, as expected, show a low degree of specialization: they are characterised by a cross sectional distribution centred around 1. For all the countries belonging to this group, the distribution function shows a quite remarkable stability over the three periods[32].

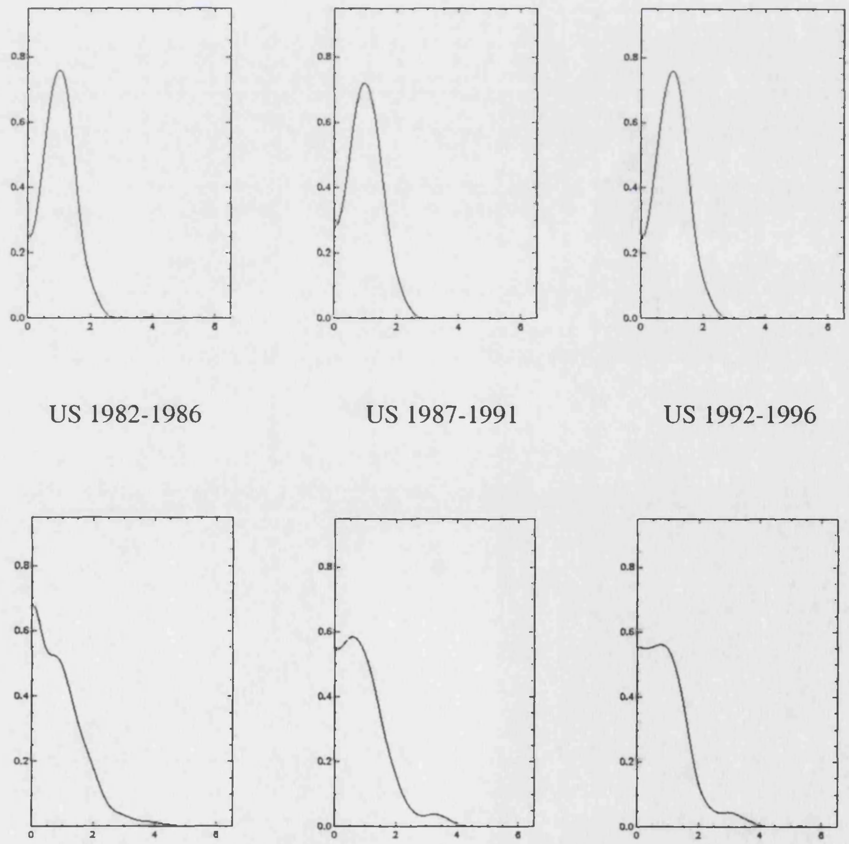
Japan is somewhat different in that it shows a much higher weight of very de-specialized technologies and a consequently larger weight of technologies with high specialization. In other words, Japan shows a higher degree of specialization than that of the above countries (as in Cantwell, 1989). This tendency appears less pronounced in the last period, when Japan experiences a

[31] The origin and destination class must be the same distance apart. For further details, see Cowell (1995).
[32] To save space, in Figure 1.3 the evolving pattern of the cross sectional *RTA* distribution is reported only for the most representative countries, and not for the other countries having a similar pattern and indicated in the text.

widening degree of technological specialization: the evolution of the cross sectional distribution is characterised by a decreasing weight of the very de-specialized technologies and a slight movement towards the right.

The remaining countries are definitely more specialized than the G5: the cross-sectional distribution of the specialization index has a declining pattern. These countries have a large number of de-specialized technologies, but also values around and above 1 appear to have significant weight (i.e. they have a long right tail).

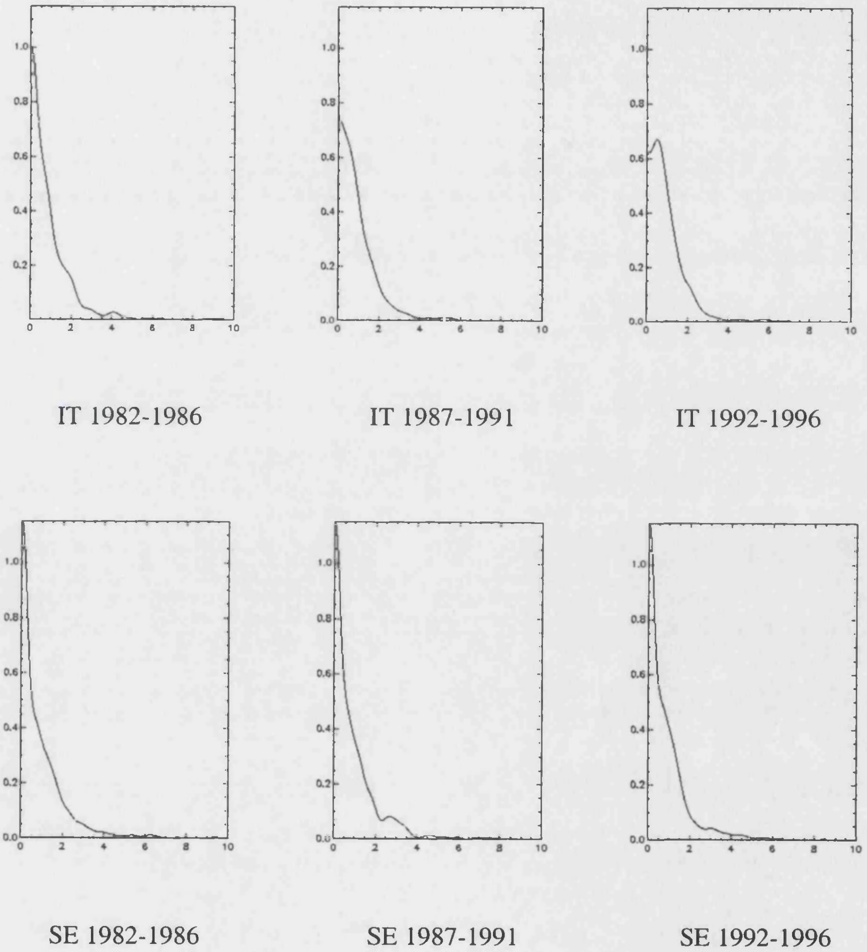
Figure 1.3 Estimated cross-sectional distributions



US 1982-1986 US 1987-1991 US 1992-1996

JP 1982-1986 JP 1987-1991 JP 1992-1996

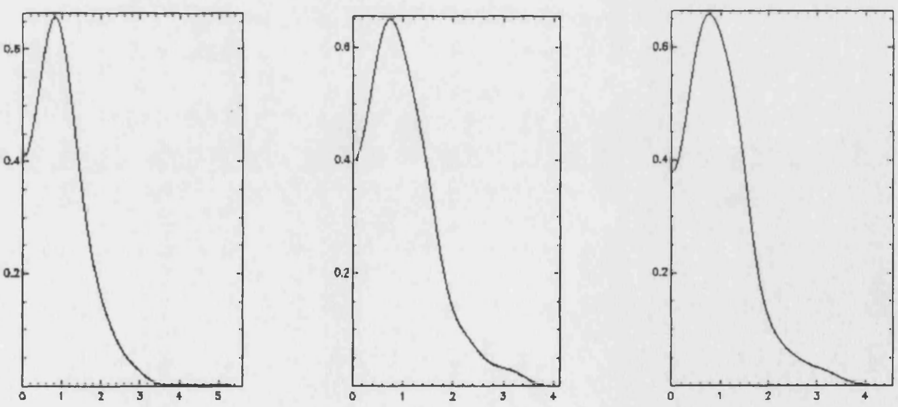
Figure 1.3 (cont.) Estimated cross-sectional distributions



All these countries, except Italy, appear quite stable with regard to the shape of the cross-sectional density function. *Italy*, however, seems to evolve [redacted] small countries towards one more similar to that of the large countries[33].

Large countries could, however, display a low degree of specialization simply because they themselves make a considerable part of the total world patenting. Since this is especially true for the US, I checked whether the US is really characterised by a low degree of specialization or this property results from the US having a very large share of patent applications at the EPO. I calculated the revealed technological advantage index, excluding US patents from the world individual technology and grand totals, and then estimated again the cross-sectional densities. The resulting cross-sectional distributions, although with heavier tails (the value of the density function at the mode is about 0.65, which is below the corresponding value in the original density function), are still centred on one, the mean value of the index, thus signalling a low overall degree of technological specialization (see Figure 1.4).

Figure 1.4 US cross-sectional distributions of the *RTA* index obtained excluding US patents from the world totals. The distributions are estimated for three sub-periods: 1982-1986, 1987-1991, 1992-1996.



[33] This finding contrasts with the apparent stability of the pattern of Italy emerging from the analysis of Laursen (2000).

1.5.2 Intra-distribution dynamics

Intra-distribution dynamics include information on switches in ranks and on the distance traversed when such switches happen. A way to quantify this phenomenon in a sequence of distributions is to obtain a Markov stochastic kernel, which gives the conditional distribution of a continuous variable at time *t+k*, given its value at time *t*. More precisely, assume that the process governing the evolution of the specialization level of a country in a technology (i.e. $RTA^{t}_{ij} = X_t$) is a *General Markov Chain*:

$$\{X_t\} : \Pr(X_{t+1} \in A_{t+1} | X_t = x, X_{t-1} \in A_{t-1}, \ldots) = \Pr(X_{t+1} \in A_{t+1} | X_t = x) \qquad (9)$$

The law of motion of the sequence of measures $\{\phi_t : t \geq 0\}$ can then be described by a first order time-invariant autoregressive process:

$$\phi_{t+1} = T^{*}(\phi_t) \qquad (10)$$

where T^* is a Stochastic Kernel and can be estimated non-parametrically from the data[34]. Hysteresis due to national external effects would imply that the probability measure ϕ_t tends towards a bimodal distribution in the long run, with very little or virtually no measurable mobility of individual field between the two modes.

For each country, two Markov stochastic kernels have been estimated to represent the conditional probability distributions of the *RTA* index for one-year and ten-year transition periods. For each cross-sectional unit (i.e. each 3-digit IPC class) a time series from 1982 to 1996 is available. The stochastic

[34] A more detailed description of Markov Chains and intra-distribution dynamics is given in Appendix B: Models for distribution dynamics.

kernels here presented are estimated considering every time series as an independent realisation of the same process and by pooling all the observations on the transitions between periods t and $t+k$ (where k is equal to 1 and 10).

The estimation procedure adopted is the following. First, an Epanechnikov kernel is used to non-parametrically estimate the joint density of the revealed technological advantage of a given country in technology j at dates t and $t+k$, choosing window width optimally, as suggested in Silverman (1986: 4.3.2). This estimated joint density implies a current period marginal density, which is calculated by numerical integration. Dividing the joint density by the estimated marginal gives the conditional density of the specialization index at time $t+k$, given the value it has at time t, i.e. the stochastic kernel graphed in Figure 1.7, which can be found at the end of the chapter. Under assumptions giving consistency of the joint density estimator, the implied marginal is also consistently estimated. Provided then that the true marginal is bounded away from zero, the stochastic kernel is consistently estimated as well[35].

For presentation, the kernels in Figure 1.7 have been drawn such that the grid lines become more finely spaced where more data were available for estimation. The graphs are obtained for ranges including the 95% of the distribution of observations (i.e. cutting off the right tail), so to avoid the problem of over-smoothing and spikes for very disperse and isolated observations. Contour plots are obtained by projecting vertically onto the floor the stochastic kernels: the contour levels have been chosen to be informative of some of the fine structure in those kernels.

Figure 1.7 shows the stochastic kernels for 1-year and 10-year transitions in the *RTA* data between 1982 and 1996[36]. Imagine cutting the one year transition stochastic kernel perpendicularly to the (*Period t*, *Period t+k*) plane, starting

---

[35] The literature on large-sample properties for density estimation is quite vast: the best reference is Silverman (1986: 3.7) and the references given there.
[36] The same remark reported in footnote 32 applies here.

from any point on the axis marked *Period t* and extending parallel to the axis marked *Period t+k.* Saying that the stochastic kernel is a conditional probability density function means precisely that the projection traced out is non-negative and integrates to unity (i.e. that projection is the equivalent of a row of a transition probability matrix, with non-negative entries summing to 1). This probability density describes transitions over *k* year(s) from a given *RTA* value in period *t* and the whole graph shows how the cross-sectional distribution at time t evolves into that at time *t+k.*

If most of the graph were concentrated along the 45-degree diagonal, then the elements of the distribution tend to remain around the values where they started from. Of course, the greater the dispersion around the diagonal, the heavier the tails of the conditional distribution, and the farther an observation can move away from its initial value, ceteris paribus. If, on the contrary, most of the mass in the graph were rotated 90 degrees counter-clockwise from that 45-degree diagonal, then substantial overtaking occurs (specialized technologies tend to become de-specialized and vice-versa). If most of the graph were concentrated around the 1-value of the *Period t+k* axis -extending parallel to the *Period t* axis- then the cross-section distribution converges towards equality to the world specialization pattern over a *k*-year horizon. More generally, if the conditional distributions appear to be the same regardless of the starting *(Period t)* value, then the stochastic kernel is one where a *k*-year transition takes any initial distribution to the same long- run cross-sectional distribution.

For the *United States,* 95 percent of the observations in the panel lie between 0 and 1.889, hence the range is quite narrow. Looking at the one-year transition stochastic kernel, technological specialization in the US appears significantly and quite equally persistent through the whole range, as most of the mass of the stochastic kernel is concentrated around the 45-degree line[37]. On the contrary, but as expected, greater dispersion signals persistence is less pronounced over a 10-year horizon.

The range containing the lower 95 percent of the *RTA* values for *Germany* is (0, 1.859), hence very similar to that of the US. Also the 1-year transition kernel is quite similar to the one for the US, suggesting a significant tendency to persistence for Germany as well. The two countries differ with reference to the 10-year transition kernel: for Germany, this is characterised by regression towards the mean for values of the specialization index above 1, an interesting feature that characterises other countries.

France and the UK show similar features. The ranges including the lower 95 percent of the observations are (0, 2.31) and (0, 2.259), respectively, hence they are wider than for the US and Germany. This is also true for the dispersion, as confirmed by the height of both the estimated kernels, which are flatter for France and, even more so, for the UK. This notwithstanding, France and the UK are characterised by persistence as the US and Germany, especially for values around the mean of the index (slightly below for the UK). There is, again, a tendency to regression towards the mean for high specialization values, more pronounced over a 10-year horizon.

The lower 95 percent percent of observations for *Japan* lies below 2.649, hence the range is of the same order of magnitude as that of UK and France, but the dispersion is somewhat lower for the 1-year transition kernel. Persistence is significant throughout the whole range, even if this time the mass

[37] Note how the conclusions about the US emerging from the cross-sectional distribution and the stochastic kernels analyses do not agree with those of Cantwell (1989), Laursen (2000) and Amendola et al. (1998). This study does not show any tendency of the US towards increasing specialization; rather this country appears as a remarkable example of stability. Most likely, the reason of the different result lies in the different data sets employed rather than in the different estimation techniques used. As mentioned earlier, the home country bias problem may adversely affect the results for the United States if US patent data are employed. This would go in favour of a greater reliability of the results obtained here in comparison with those of the earlier studies.

is concentrated just above the diagonal. The 10-year transition kernel is, again, characterised by relatively strong persistence and regression towards the mean for values above 1.5.

For the remaining countries, the upper limit of the lower 95 percent range lies between 2.908 (Switzerland) and 4.169 (Austria); hence, the ranges are wider than those of the first group of countries. The 1-year and 10-year transition kernels of four countries (*Switzerland, the Netherlands, Italy, and Sweden*) have common characteristics. The 1-year kernel's pick follows the 45-degree line, remaining right above it, except for very low values. In other words, over one year, the specialization index shows some sign of persistence around (possibly slightly below) its initial value, when this is not too low. When it is, there is, instead, a tendency to move a little to the right (i.e. to increase), but not very far. This tendency is confirmed in the 10-year transition kernels, while medium and high values are characterised by regression towards the mean. This phenomenon appears stronger for Sweden and Switzerland than for Italy and the Netherlands[38]. Note that for Italy, over ten years, there's a higher probability of very disadvantaged technologies to improve specialization: at very low values the kernel is much more centred around 0.5 on the *Period t+10* axis[39]. Note also that Italy shows in the 10-year transition kernel a dispersion similar to (and not higher than) that in the 1-year transition (the height of the two kernels is approximately the same). Both these findings are consistent with the different evolution of Italy's cross-sectional distribution, shown in the previous section.

---

[38] *Austria* is somewhat particular in that the 1-year transition kernel is similar to the 10-year transition one; i.e. it already shows strong regression towards the mean from above.

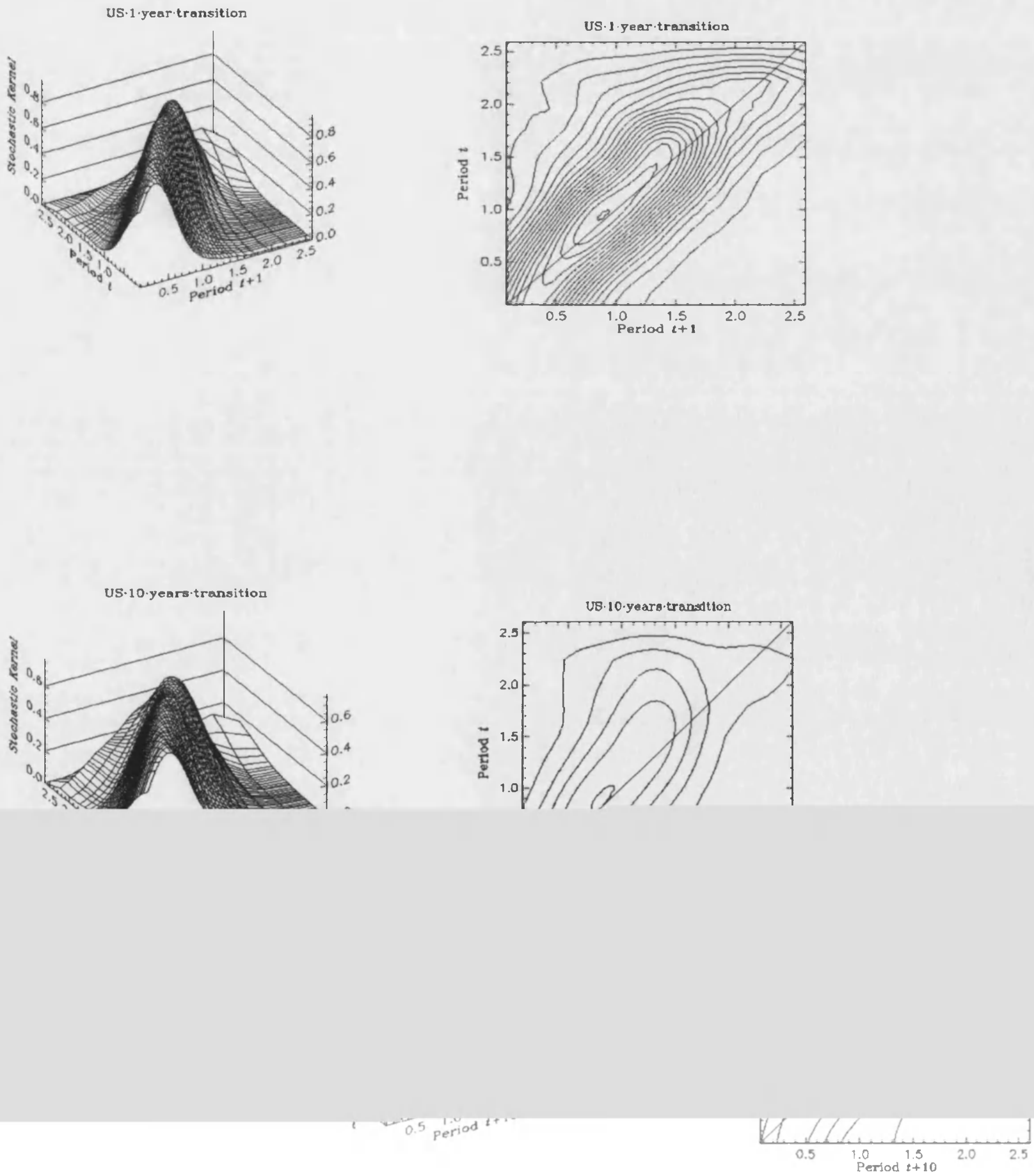
[39] The technologies where Italy significantly improved its specialization, moving from a state of de-specialization (*RTA*<0.5) to one of relatively high specialization (*RTA*>1,5) include: B06 (generating and transmitting mechanical vibrations), A63 (sports, games, amusements), B32 (layered products), A22 (meat treatment, processing poultry or fish), B65 (packing, storing), E21 (earth drilling, mining), C11 (animal or vegetable oils; detergents; candles), C09 (dyes, paints), G10 (musical instruments, acoustics), D06 (treatment of textiles), B04 (centrifugal apparatus), H04 (electric communication technique).

Once again, the stochastic kernels for the US have been estimated excluding US patents from the totals to check for the robustness of the lower mobility (higher persistence) result for large countries. As expected, the estimated stochastic kernels are characterised by higher dispersion. Nevertheless, the one-year transition stochastic kernel is still centred around the main diagonal and so is the ten-year transition one for the range of values of the *RTA* index above 0.5 and below 1.5, that is in the range where most of the values are observed. Outside this range, the shape of the estimated stochastic kernel is consistent with the asymmetry result.

The estimation of Markov stochastic kernels provides insightful visual evidence, but leaves unsolved the fundamental problem of evaluating the extent of mobility or persistence. A way towards solving this problem is to interpret the operator T^* as a transition probability matrix. This is done discretising the state space of specialization index values, that is dividing it into discrete cells that span the space of all possible realisations (the interval [0,N]). The resulting transition probability matrix describes the conditional probability of transitions between cells and can be easily estimated by Maximum Likelihood, counting the observed transitions out of each discrete cell into itself or the other cells, and then normalizing this count by the total number of observations starting from that particular cell (Basawa and Prakasa Rao 1980).

The reasons why the transition probability matrix represents a very useful tool are twofold: on the one hand, the transition probability matrix is easy to interpret; on the other hand it can be used to quantify mobility, perform cross-country comparisons and obtain the long run stationary distribution, where it exists.

Figure 1.5 Stochastic kernels and contour plots for the US, obtained excluding US patents from the world totals.



of
his
se,
ear

Table 1.3 (at the end of the chapter) presents estimates of the probability transiting between different grid cells of a country's distribution of *RTA*. T is done for both one-year and ten-year transition periods. In each ca boundaries between the cells have been chosen such that technology-ye observations are divided roughly equally between the grid cells. As a consequence, each grid cell corresponds approximately to a quantile of the *RTA* distribution across technologies and over time. The values of estimated transition probabilities characterise the degree of mobility between different quantiles of this distribution.

Each panel in the table has to be interpreted as follows. The numbers in the first column are the total numbers of technology-year observations beginning in a particular cell, while the first row of numbers denotes the upper point of the corresponding grid cell[40]. Row j includes the estimated probability of remaining in state j (i.e. the element (j,j) of the matrix) and of moving from state j to state s. Estimated values of transition probabilities close to one along the diagonal indicate persistence in a country's pattern of *RTA*, while large off-diagonal terms imply greater mobility. The results of Table 1.3 for the one-year transition period suggest a significantly high degree of mobility in patterns of international technological specialization in all the countries: the probability of moving out of the original state ranges from 28 percent to 70 percent. Even for the US all the intermediate grid cells along the main diagonal have values below 0.5, and this is almost always true for all the other countries, excluding Japan. This finding is consistent with the visual evidence presented earlier: the stochastic kernels were indeed all characterised by quite a high spread. Note also that movements out of the intermediate grid cells are directed towards the neighbour upper and lower cells in equal proportions: observations show the tendency to wander about in all directions.

Mobility is somewhat lower at the bottom of the distribution: it is more difficult for a country to improve its level of specialization in those technologies where it is very disadvantaged[41]. The difficulty to move out of a

---

[40] Note how for Austria the first cell is not an interval, but a single point at zero. This happens because of the great number of zero observations, a clear sign of high degree of specialization

[41] Mobility appears lower also at the top, even if to a lesser extent compared to what happens at

situation of strong relative disadvantage signals the importance of own research for the ability of a country to innovate. A country that today does not have enough experience (i.e. accumulated knowledge) and capabilities in a technology area to be an innovator will find it increasingly difficult to innovate in the future. Even if R&D output can be transferred intentionally (through international patents and licenses) or unintentionally (through spillovers), it is not necessarily adopted and further improved in a country unless this country has developed itself some previous knowledge in the technological area, i.e. unless the country has reached some *threshold* level of knowledge in the area[42].

For the one-year transitions, the final row of each panel gives the implied ergodic distribution, that is the asymptotic unconditional probability of being in each state (i.e. the probability of being in each state regardless of the initial state). It is obtained iterating the estimated transition probability matrix forward in time and allowing the number of iterations to tend towards infinity[43]. If the estimated fractile matrix is ergodic, its stationary distribution will be uniform relative to the quantiles (Quah, 1993b), as indeed is the case for all the estimated matrices. This means that even if countries were increasingly specializing in a subset of technologies, one could not observe a polarization of the ergodic distribution towards extreme values and the emergence of a bimodal distribution.

A way to check if this tendency is indeed at work is to examine whether the upper and lower quantiles move towards the extremes or, alternatively, whether

the bottom of the distribution. Note, however, that the grid cell at the top is defined by a very wide range, whereas the others are not. Given the results of the stochastic kernel analysis, it should be it clear that this could hide possibly relevant dynamics. For this reason, it is preferable not to draw conclusions from the apparently high persistence in the top state.

[42] Cohen and Levinthal (1990) first recognized the ability to exploit external knowledge as a critical component of innovative ability and named it "absorptive capacity".

[43] Note that for the stability of the transition probability matrix and the existence of the ergodic distribution the highest eigenvalue has to be equal to 1 and all the others need to be smaller than 1 in absolute value. This happens for all the countries, the first eigenvalue being equal to 1 within a two-digit approximation.

the interquartile range increases over time. This can be done by running a simple linear regression of the interquartile range on time and testing for the significance of the slope coefficient. A significant positive time trend in the interquartile range combined with high persistence (i.e. large entries on the main diagonal - especially at the ends - of the estimated fractile matrix) would point to polarization of the distribution, whereas a negative or no time trend in the interquartile range and low persistence in the transition matrix would not support that hypothesis. All country regressions of the interquartile range on time yield zero or negative slope coefficients, thus confirming the absence of any tendency towards polarization and deepening of initial specialization patterns[44].

The second half of Table 2 shows that, as expected, the degree of mobility is higher over a ten-year transition period. Still, the same tendencies that characterise the one-year dynamics appear also here. The results for Italy are peculiar in that they show a probability of transition out of the bottom grid cell much higher than that of the other countries. This is consistent with the findings on the changing shape of the cross-sectional distribution in the previous section[45].

Table 1.1 calculates a variety of mobility indices (following Shorrocks 1978; Geweke et al. 1986; Quah 1996b) for each of the countries. Each of these indices attempts to reduce information about mobility from the matrix of stationary one-year transition probabilities M to a single statistic. Thus, μ_1 evaluates the trace of the matrix $(tr(M))$, μ_4 analyses the determinant $(det(M))$, and μ_3 and μ_5 are based on the eigenvalues λ_j of the matrix. Finally, μ_2 presents

[44] There is actually one exception: Germany. In this case, the slope coefficient is significantly positive, but very low (0.01) and is originated by a widening of the interquartile range in the nineties. Once again, this is probably the consequence of unification between West and East Germany.

[45] Cefis and Orsenigo (2001) find that Italy appears to be less persistent than the G5 countries; this analysis further shows that this country has been characterised by a probability of moving out of the de-specialization state higher than the other countries and increasing in time.

information on the average number of class boundaries crossed by an observation originally in state k, weighted by the corresponding proportion π_k of the ergodic distribution.

Table 1.1 Some mobility indices.

	μ_1	μ_2	μ_3	μ_4	μ_5
US	0.598	0.665	0.598	0.987	0.292
Germany	0.645	0.740	0.645	0.994	0.321
France	0.710	0.906	0.710	0.998	0.474
UK	0.770	1.012	0.770	0.999	0.556
Japan	0.545	0.574	0.545	0.974	0.232
Italy	0.700	0.869	0.700	0.997	0.431
Switzerland	0.665	0.824	0.665	0.993	0.416
Netherlands	0.645	0.808	0.645	0.992	0.393
Sweden	0.648	0.867	0.648	0.996	0.410
Austria	0.668	0.930	0.667	1.000	0.403

$\mu_1 = \frac{n - tr(M)}{n-1}$; $\mu_2 = \sum_k \pi_k \sum_l p_{kl}|k-l|$; $\mu_3 = \frac{n - \sum_j |\lambda_j|}{n-1}$; $\mu_4 = 1 - |\det(M)|$; $\mu_5 = 1 - |\lambda_2|$

The results in Table 1.1 confirm that the overall degree of mobility is quite high. Among the countries in the sample, Japan has the lowest value of all the indices: its technological specialization appears as the least subject to shifts from one year to the next, as measured by the cross boundaries transitions under the stationarity assumption. The US comes right after and seems to be somewhat equally distant from Japan and Germany, which follows. Moving towards higher values of the mobility indices, there is then a group of three countries very close to each other (the Netherlands, Sweden and Switzerland), followed by Italy and then, surprisingly, France and the UK. The evidence on Austria is rather mixed.

Table 1.3 does not seem to reveal any particular striking difference between different countries. In order to test for the statistical significance of the similarities between the dynamics of different countries, the asymptotic properties of first-order Markov chains derived in Anderson and Goodman (1957) can be used. The two authors show that, for each state k, under the null hypothesis $p_{kl} = \tilde{p}_{kl}$,

$$\sum_{l=1}^{n} m_k^* \frac{(p_{kl} - \tilde{p}_{kl})^2}{\tilde{p}_{kl}} \sim \chi^2(n-1), \quad m_k^* \equiv \sum_{t=0}^{T-1} m_k(t) \tag{11}$$

where p_{kl} is the estimated transition probability from state k to state l, $\tilde{p}_{kl}$ is the corresponding probability of transition under the known null, and $m_k(t)$ denotes the number of technologies in cell k at time t.

The test statistic above cannot be used to directly test the hypothesis that, for each state k, the transition probabilities estimated for any two countries are the same. This is because both sets of transition probabilities are estimates, hence the null hypothesis cannot be properly formulated. However, following Proudman and Redding (2000), I shall adopt the null hypothesis that the Data Generating Process (DGP) underlying the pattern of *RTA* of, say, country A is the estimated transition probability matrix of country B ($p_{kl}^A = \tilde{p}_{kl}^B$). It is then possible to test whether the transition probabilities estimated for country A are significantly different from those of the null. Similarly, one may then test whether the estimated transition probabilities for country B are significantly different from the null that the DGP is country A's matrix of transition probabilities ($p_{kl}^B = \tilde{p}_{kl}^A$). These tests may be undertaken for each state $k=1,\ldots,n$. Furthermore, since the transition probabilities are independently distributed across states k, it is possible to sum over states and test the hypothesis that for all states $k=1,\ldots,n$, the estimated transition probabilities are equal to those under the null. The resulting test statistic is asymptotically distributed as a $\chi^2(n(n-1))$.

The test has been implemented for each pair of countries in both directions (i.e. using alternatively the transition probability matrix of each of the country in the pair as the null). The null is almost invariantly rejected[46], that is idiosyncratic elements are quite strong and do affect the dynamics of countries' specialization patterns. These specificities could originate from the institutions and mechanisms supporting technological innovation, which might greatly differ among countries, and from other elements like factor endowments.

1.6 Concluding Remarks

The results emerging from the econometric analysis clearly emphasise the existence of two strong *country size effects*: one static and one dynamic. The first one is well known and has been accounted for in previous analyses[47]: economically large countries are less specialized and spread their innovation activities across a wider range of technologies. The analysis on the shape of the specialization index distribution has shown that this is fairly symmetric around one for the most industrialized countries, with the partial exception of Japan whose distribution is more skewed to the right, but less and less so in time. This static size effect is also confirmed by a simple cross-country linear regression of the degree of specialization, as measured by the Gini coefficient, on the manufacturing output, a proxy for size (Table 1.2)[48].

[46] The only exception is the Italy-Switzerland pair, for which the two hypotheses specified above cannot be rejected at the 95% confidence level.
[47] See, for example, Archibugi and Pianta (1992b).
[48] The regressions reported in Table 3 are performed on a relatively small sample size. However, the results are confirmed and even reinforced when the sample size is extended to include other countries with fewer patent applications at the EPO than the countries analyzed here. Regressions (1) and (2) are also relative to the initial year of the sample: the same results

There is then a second size effect, which could not emerge from previous studies. Because the approaches adopted in those studies are fundamentally static in nature and are averaging across observations in various ways, they have difficulties in capturing the genuine dynamic forces characterising the evolution of countries' technological specialization patterns. The distribution dynamics analysis performed in this study has shown that economically large countries are also characterised by a higher degree of persistence (i.e. lower degree of mobility), this result being less strong for France and the UK[49]. This means that the specialization level of large countries in individual technologies displays lower probability of moving around and far from its initial level. Again, this result is confirmed by a simple linear regression of a mobility index on the size of a country (see Table 1.2, column (2)).

Regardless of the distinction between "large" and "small" countries, the tendency towards persistence is never pronounced: technological specialization in advanced countries displays fluctuations around and far from its initial level with a probability almost always higher than 0.5[50]. Furthermore, mobility in technology appears higher than that emerging from trade analyses[51], thus weakening the case for causality from hysteresis in technological development to hysteresis in trade specialization patterns[52]. This result seems to undermine the theory of technological accumulation: the apparent stability of technological specialization patterns, as represented by the cross-sectional distribution of the *RTA* index, hides a significant amount of mobility even in

are obtained when the regressions are performed, for example on the cross-section of countries in any other year of the sample period.

[49] Recall the results on the mobility indices for France and the UK, which rank them as the most mobile countries among the ten analysed in detail.

[50] Mobility is also invariantly higher over ten than over one year transition period (Italy being the only exception among the countries here examined), a result consistent with those of Cefis and Orsenigo (2001), who find that persistence in firms' patenting activity declines significantly as the transition period lengthens.

[51] See, again, Proudman and Redding (1998 and 2000) and Brasili et al. (2000).

[52] Recall that localised spillovers should have their most direct effect on technological specialization.

the most industrialised and large countries.

Table 1.2 Some simple cross-sectional regressions

	(1)	(2)	(3)
Variable	Gini	M2	Gini1996/Gini1982
C	0.411**	0.916**	1.255**
	(7.403)	(19.22)	(10.409)
Isic300_US	-0.28*	-0.328**	
	(-2.035)	(-2.77)	
Gini1982			-0.745*
			(-2.25)

Note:

(1) Regression for the static size effect. The Gini coefficient is regressed on a constant and each country's manufacturing output (ISIC 300) measured in dollars and relative to US.

(2) Regression for the dynamic size effect. The M2 mobility index (which also measures mobility outside the main diagonal) is regressed on a constant and each country's manufacturing output (ISIC 300) measured in dollars and relative to US (Isic300_US).

(3) Regression of the change in the Gini coefficient over the sample period on a constant and its initial value.

* Significant at the 10 percent level. ** Significant at the 5 percent level.

Data on manufacturing output and the exchange rate are taken from the OECD STAN database (note that manufacturing output is not available for Switzerland, which had to be excluded from the sample when that variable was involved).

Mobility is also *asymmetric*: it seems to be mostly difficult for a country to improve specialization in technologies where it is greatly disadvantaged, while high specialization shows a fairly general tendency to revert towards lower levels. Indeed, from both the estimated stochastic kernels and transition probability matrices it emerges that observations tend to revert towards the mean, but only from one side of the distribution. Furthermore, on average, but with the exception of Italy, the probability of remaining a highly specialized country declines more than that of remaining an occasional innovator as time goes by. This clearly explains the decline in persistence (i.e. the increased mobility) as the transition period lengthens.

Asymmetry in technology dynamics suggests that even if R&D spillovers were international in scope, countries need to have some prior level of knowledge, R&D investment, or complementary assets in the relevant technology to be able to understand and employ knowledge produced elsewhere. In the absence of a sufficiently high *absorptive capacity* (Cohen and Levinthal 1990), originated from previous experience in a technology, countries are not likely to overcome their weaknesses. Since the speed of reversion to the mean from above is inversely related to country size, this result, again, does not support the existence of self-enforcing mechanisms deepening initial specialization patterns, or even locking them in. If that were the case, not only a country should retain its initial comparative advantage in a field and possibly reinforce it, but also this persistence effect should be stronger the more the country's initial specialization pattern is skewed.

In sum, the empirical dynamics of technological specialization emerging from the analysis of industrial countries does not seem to support the idea that there are cumulative and reinforcing mechanisms at work, which could then generate path-dependence in the original technology and trade specialization patterns. If there is persistence in the trade patterns of industrial countries, this could then be the consequence of these countries occupying a relatively stable position in the international economy in terms of factor endowments. Mobility in trade patterns could instead be the consequence of cross-country mobility of technology, which would be consistent with the observed changes in patterns of technological specialization. Furthermore, the immediate normative implication of the results outlined above is that targeted industrial and technology policies might not be effective, because an initial comparative advantage can be eroded by the knowledge on which it is based flowing to foreign competitors. There is, however, one notable exception: policies aimed at building competitive ability in very disadvantaged technologies may take a country out of an otherwise enduring weak position in the international arena.

Appendix A: The International Patent Classification

The International Patent Classification (IPC)[53] provides a common classification for patents and published patent applications. It is a hierarchical classification primarily concerned with the technological characteristics of the innovation. It is designed to represent the whole body of knowledge, which may be regarded as proper to the field of the invention and it is based on both an "application principle" and a "functional principle".

The classification attempts to ensure that any technical subject, with which an invention is essentially concerned, can be classified, as far as possible, as a whole and not by separate classification of constituent parts. The technical subjects refer to either the intrinsic nature of the invention, or its function or the way it is used or applied, while an invention can be either a product or a process.

A patent is assigned to an IPC class according to the following general guidelines: (i) if the object of the patent has a very specific product application, then it is classified into a technology class according to the application principle; (ii) if the innovation has a broader field of application, then the patent is assigned to multiple classes according to both the application and the functional principles; (iii) finally, if no dominant field of application exists, then the patent is assigned to a class which corresponds to the function the innovation is aimed at fulfilling, hence according to the functional principle.

The IPC is divided into eight sections:

A Human Necessities

B Performing Operations; Transporting

C Chemistry; Metallurgy

[53] The IPC was established with the Strasbourg Agreement in 1971 and entered into force on October 7, 1975.

D Textiles; Paper

E Fixed Constructions

F Mechanical Engineering; Lighting; Heating; Weapons; Blasting

G Physics

H Electricity

Each section is itself divided into classes, sub-classes, groups and sub-groups, in descending order of hierarchy. Solely the number of dots preceding their titles determines the hierarchy among sub-groups. For example:

Section:	B	Transporting
Class:	B 64	Aircraft, Aviation, Cosmonautics
Sub-class	B 64 C	Aeroplanes, Helicopters
Main group	B 64 C3 /00	Wings
One-dot sub-group	B 64 C3 /10 •	Shape of wings
Two-dot sub-group	B 64 C3 /14 • •	Frontal aspect
Etc.		

In the chapter, the unit of analysis is a 3-digit technology, i.e. a class in the IPC hierarchy represented above.

Appendix B: Models for distribution dynamics[54]

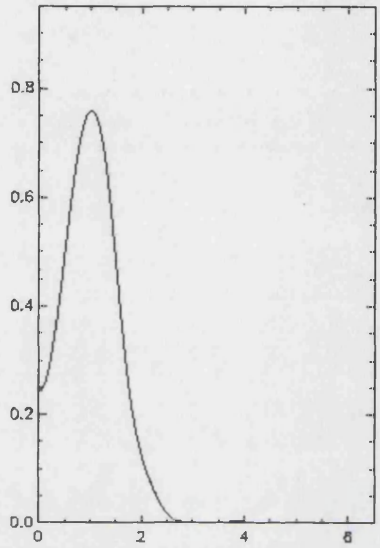
This appendix briefly describes the statistical methodology employed in the econometric analysis of the chapter. The empirical model adopted is that of *Distribution Dynamics*, recently employed in cross-country growth literature (Quah 1996, 1996a, 1996b, 1996c, 1997). The purpose of distribution dynamics is to study how the distribution of a certain economic variable evolves in time. In the present chapter, the variable of interest is the revealed technological advantage index (RTA). For a country, this index is calculated in each of a given set of technologies and takes a positive real value in each of them. So if there are N technologies, there are also N values of the index for each country. Hence, what the interest here is in the distribution of the index across technologies for a country.

Consider such a distribution: Figure 1.6 shows an empirical density function of the *RTA* index calculated for the US for a representative year. The values on the horizontal axis represent the value of the index in individual technologies. The evolution over time of the distribution of the *RTA* index is represented by changes in its external shape and by intra-distribution movements. The mechanism governing this evolution can be summarised by a transition function, which maps each point in the distribution from one period to the next and describes a Markov process.

The approach of distribution dynamics differs from the traditional Markov process theory approach, where the emphasis is on a scalar process, from which an unobservable sequence of probability distribution is inferred. Here, instead, a sequence of entire (empirical) cross-sectional distributions is actually observed, while the (*dual*) scalar process is implied, but never observed (Quah, 1996c).

[54] This section is based on various papers by Quah (1996a, 1996b, 1996c, 1997), on Silverman (1986) and Basawa and Prakasa Rao (1980).

Figure 1.6 An empirical cross-section distribution.



B.1 Markov chains

Call the variable of interest, at time t, X_t (with t an integer) and assume it can take values in a set E. The process $\{X_t, t > 0\}$, can be defined as a Markov chain if, given the present state x, the probability of transition to a certain state in the set A_{t+1} in the next period only depends on the current state of the process, and is independent from the past:

$$\Pr(X_{t+1} \in A_{t+1} | X_t = x, X_{t-1} \in A_{t-1}, \ldots) = \Pr(X_{t+1} \in A_{t+1} | X_t = x)$$

where the A_t is a subset of E.

the probability that the next step will take the process in a certain set A, given that the current state is x:

$$P(x, A) = \Pr(X_{t+1} \in A | X_t = x)$$

for all values of x in E and all the subsets A.

Now let $p(x, y)$ be a measurable function for which:

$$p(x, y) \geq 0 \quad \text{and} \quad \int_E p(x, y) dy = 1$$

where x and y are points in E. Suppose that the kernel $P(x, A)$ can be defined as the integral of this function over the set A:

$$P(x, A) = \int_A p(x, y) dy$$

then $p(x, y)$ is the *transition density function* associated with $P(x, A)$.

When X_t is discrete, that is it can assume only a finite or countable number of values[55], the process $\{X_t, t > 0\}$ is described by a *transition probability matrix*, i.e. a square array of non-negative numbers with row sums equal to 1:

$$M = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ & & \cdots & \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

[55] i.e. The set E is either finite or countable infinite.

where:

$$p_{ij} = \Pr\left(X_{t+1} = x_j \middle| X_t = x_i\right)$$

$$\sum_{j=1}^{n} p_{ij} = 1 : \ p_{ij} \geq 0$$

That is, the matrix element in position (i, j) denotes the conditional probability of a transition to state j at time t+1, given that the system is in state i at time t. Each row of the matrix then denotes the conditional probability distribution at time t+1 of an individual in state i at time t.

A transition probability matrix is said to be *stationary* or homogeneous when the conditional probabilities depend on the time interval of transition, but not on the time t. For such a chain s-step transition probabilities can be defined:

$$p_{ij}^{(s)} = \Pr\left(X_{t+s} = x_j \middle| X_t = x_i\right) = \Pr\left(X_s = x_j \middle| X_0 = x_i\right)$$

By a recurrence argument:

$$(p_{ij}^{(s)}) = M^s$$

One question that arises is whether after a sufficiently long period of time the system settles down to a condition of statistical equilibrium in which the state occupation probabilities are independent of the initial conditions. If this is the case, the limit π is called *ergodic distribution.* It must satisfy and be the unique solution of the equation:

$$\pi' = \pi' M : \pi' \cdot 1 = 1$$

B.2 Distribution dynamics

Let F_t be the distribution of X_t at time *t*. Associated with it is a probability measure ϕ_t. Describe the law of motion of the sequence of measures $\{\phi_t : t \geq 0\}$ by a first order autoregressive process[56] (Quah, 1997):

$$\phi_{t+1} = T_t(\phi_t)$$

where the operator T_t maps the distribution from period *t* to period *t+1*. In the chapter I make the assumption that T_t is time invariant, so that it is possible to write:

$$\phi_{t+1} = T^*(\phi_t)$$

The characteristics of T^* depend crucially on those of the variable X_t. If X_t is discrete the operator T^* can be interpreted as the stationary transition probability matrix *M* of a Markov process:

$$\phi_{t+1} = M'\phi_t$$

However, if X_t can take continuous values, for example any value on the real line, then the operator T^* must be interpreted as a transition function or stochastic kernel $P(x,\cdot)$. The distribution at time t+1 is then defined by:

[56] In general, ϕ_t might show more than first-order dependence. In that case, the equation describing the law of motion should be modified to permit that.

$$\phi_{t+1}(A) = \int P(x, A)\phi_t(dx)$$

where A is a subset of E.

The transition function P must satisfy the following two properties:

(i) for every set A in E, the function $P(\cdot, A)$ is measurable;

(ii) for every point x in R, the function $P(x,\cdot)$ is a probability measure.

B.3 Discretisation

The stochastic kernel is a useful tool to analyse the dynamics of the entire distribution of a process. There are some reasons why, however, it may be useful to "discretise" the state space, that is to partition the continuous state space in a finite number of intervals. These sets would then constitute the states of a newly defined finite Markov process. The reasons why this approach is appealing are numerous. First, the theory of finite state space Markov processes is accessible and well developed; then, the estimation of the transition matrix is computationally simpler, results are easier to interpret and present, and many indices and statistics can be easily computed.

An arbitrary discretisation corresponds to creating a partition of the space into a finite number of subsets $A_1, \ldots\ldots, A_J$ and then associating each subset with a distinct state in a discrete state space. This is equivalent to creating a sequence:

$$\eta^{(t)} = \sum_{j=1}^{J} j\mathrm{I}_{A_j}(X_t)$$

where $\mathrm{I}_{A_j}(X_t)$ is the indicator function:

$$\mathrm{I}_{A_j} = \begin{cases} 1 & \text{if } X_t \in A_j \\ 0 & \text{otherwise} \end{cases}$$

The sequence $\eta^{(t)}$ is then treated as a discrete Markov chain[57].

In the empirical analysis I use a discretisation method thanks to which a *fractile*[58] *transition probability matrix, M*, is obtained from a stochastic kernel (Quah, 1997). The method can be easily described by providing a constructive definition of M in this case.

Fix a positive integer n: this will be the number of cells in the discretisation of the basic data. Then represent the stochastic kernel P_t[59] by the pair $(M(t), q(t))$, where $M(t)$ is the fractile transition matrix and $q(t)$ is an *n-element quantile set*, i.e. a collection of n disjoint random intervals. Denote the basic data by

$$\{x_j(t) : j = 1, 2, \ldots, N\,;\; t = 0, 1, \ldots, T\}$$

where j denotes the cross-sectional units and t indexes time. The sequence ϕ_t relates to the basic data by

$$\forall r \in R: \;\; \phi_t((-\infty, r]) = \#\{j : x_j(t) \leq r\} \times N^{-1}$$

[57] It should be mentioned that, when underlying observations are, as in this case, continuous variables, such a discretisation could distort dynamics in possibly important ways. The most extreme consequence of this would be a failure of the fundamental Markov property that the state occupied by the system in period t depends only on the state the system occupied in period $t-1$ and not on the previous one. For an extended discussion on the topic and the description of a robust discretisation method see Bulli (2001).

[58] A transition probability matrix is said to be a *fractile* when it describes transitions out of cells containing equal fractions of the entire distribution.

[59] Here I consider the most general situation in which the stochastic kernel is time dependent. When, instead, it is time invariant, then also M and q will be so.

Every fixed positive integer n implies a unique set of equally spaced probabilities:

$$\{m/n : m = 0, 1, \ldots, n\}$$

Define at time t the quantiles

$$(quant)_m(t) = \inf \{r \in R | \phi_t((-\infty, r]) > m/n\} \qquad m = 1, 2, \ldots, n$$

and take

$$(quant)_0(t) = -\infty .$$

These give the consecutive disjoint random intervals:

$$q_m(t) = ((quant)_{m-1}(t), (quant)_m(t)] \qquad m = 1, 2, \ldots, n$$

which, in turn, comprise the quantile set

$$q(t) = \{q_m(t) \; : \; m = 1, 2, \ldots, n\} .$$

By construction, $\phi_t(q_1(t)) = \phi_t(q_m(t))$ for all m, i.e. the elements of every quantile set have equal measure.

The sequence of quantile sets together with the basic data defines the transition probabilities in $M(t)$, whose (l, m) entry is:

$$M_{lm}(t) = \frac{\#\{j : x_j(t+1) \in q_m(t+1) \text{ and } x_j(t) \in q_l(t)\}}{\phi_t(q_l(t))} \qquad l, m = 1, 2, \dots, n$$

The obtained *M(t)*, which comprises all non-negative entries and has row sums equal to 1, is fractile, i.e.

$$\left(\sum_{m=1}^{n} M_{lm}(t)\right)\phi_t(q_l(t)) = \phi_t(q_l(t)) = \phi_t(q_1(t))$$

is the same for all *l*.

B.4 Mobility indices

With the discretisation of a stochastic kernel into a fractile, *M* encodes information on mobility, while *q* encodes information on shape. *M*'s role can be further clarified using mobility indices (Geweke et al., 1986; Quah, 1996; Shorrocks, 1978). Analogous to measures of income inequality – summarising the information in an entire distribution into a single scalar – a mobility index collapses into one number the mobility information in a transition probability matrix. However, just as for inequality measures, no single mobility index need be completely satisfactory. Thus I consider five of them.

First take Shorrock's index μ_1 defined by:

$$\mu_1(M) = \frac{n - tr(M)}{n-1} = \left(\frac{n}{n-1}\right)\left\{n^{-1}\sum_j (1 - M_{jj})\right\}$$

where M_{jj} denotes the *j*-th diagonal entry of the matrix *M*. Since $(1 - M_{jj})$ is the probability of exiting state *j* and $1/(1 - M_{jj})$ is the mean exit time from state j (or the average length of stay in state j) then μ_1 is the inverse of the harmonic mean of expected durations of remaining in a given part of the cross-section distribution (state j), normalised by the factor $n/(n-1)$. It thus provides one natural index of mobility: the higher is μ_1, the less "persistence" is there in M.

Since the trace of a matrix equals the sum of its eigenvalues, Shorrock's index can also be written as:

$$\mu_1(M) = \frac{n - \sum_j \lambda_j}{n-1}$$

where λ_j are the eigenvalues of M. Thus when M's eigenvalues are all real and non-negative, Shorrock's μ_1 is identical to another index I consider (Quah, 1996 and 1997):

$$\mu_3(M) = \frac{n - \sum_j |\lambda_j|}{n-1}$$

In general, however, estimated μ_1 and μ_3 will differ.

To see the motivation behind μ_3 recall that every stochastic matrix M always has one eigenvalue equal to unity, and all its other eigenvalues bounded from above by 1 in modulus. In the most regular case, when M implies a unique ergodic distribution, the sequence $\{M^k : k \geq 1\}$ converges to that distinguished matrix having all rows equal to the ergodic distribution[60]. Convergence occurs at a geometric rate, given by the powers of the

[60] When M is a fractile, the ergodic distribution is always uniform.

eigenvalues λ_j. Thus the smaller is the modulus of an eigenvalue, the larger is $1-|\lambda_j|$, the faster does the corresponding component in M^k converge. Hence, μ_3 relates positively to the average rate of convergence of the cross-section distribution towards the ergodic limit, thus sensibly indexing mobility.

When all eigenvalues except the unit one are strictly less then one in modulus, then as k grows, the dominant convergence term is given by $|\lambda_2|$, the modulus of the second largest eigenvalue.

Thus for the same reason for which μ_3 is sensible, one might also consider the following as an index of mobility (Geweke et al., 1986; Quah, 1996 and 1997):

$$\mu_5(M)=1-|\lambda_2|$$

Like μ_3, this indexes the speed of convergence. But whereas μ_3 incorporates all the different rates of convergence, μ_5 captures only the asymptotic rate. The two would be identical (up to a scaling involving only n) when evaluated at an M, whose smallest eigenvalues, beyond the largest two, turn out to be zero.

A final index related to the eigenvalues of the transition probability matrix and suggested by Shorroks (1978) is:

$$\mu_4(M)=1-|\det M|$$

Finally, Shorroks (1978) discusses yet another measure of mobility:

$$\mu_2(M)=\sum_k \pi_k \sum_l p_{kl}|k-l|$$

The summation over l gives the average number of state boundaries crossed by an individual originally in state k, and these are then weighted by the proportions in the corresponding equilibrium (ergodic) distribution π_k.

B.5 Estimation

In the empirical analysis of the chapter I use non-parametric estimation techniques to obtain empirical cross-sectional distributions of a single variable of interest at a point in time and bivariate distributions of that same variable at two different time periods. From this joint distribution one can then estimate the cross-sectional distribution of the variable of interest at the more recent time period, conditional on its starting value (i.e. the value it had in the more remote time period). This section briefly reviews the kernel method for univariate and bivariate density estimation, which will be used in the empirical analyses; for further details the reader is referred to Silverman (1986). At the end of this section, the Maximum Likelihood Estimator used to obtain the transition probability matrices is also derived.

Univariate density estimation

Assume we have a sample $X_1, \ldots, X_N$ of independent, identically distributed observations from a continuous univariate distribution with probability density function f, which we are trying to estimate. The *kernel estimator* with kernel K is defined by:

$$\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - X_i}{h} \right)$$

where h is the *window width* (also called *smoothing parameter* or *bandwidth*) and the *kernel function K* satisfies the condition[61]:

$$\int_{-\infty}^{\infty} K(x)dx = 1$$

Provided the kernel K is everywhere non-negative and satisfies this condition (i.e. it is a probability density function), $\hat{f}$ will itself be a probability density function and will inherit all the continuity and differentiability properties of the kernel K.

The problem of choosing how much to smooth is of crucial importance in density estimation. There is no universally accepted approach to this problem. However, Silverman (1986) gives automatic choices with optimal properties. For example, the optimal window width with a Gaussian kernel[62] is

$$h = 0.9AN^{-1/5}$$

where A is an adaptive measure of spread equal to

A = min (standard deviation, interquartile range/1.34)

In my case, I here deal with positive domains, i.e. domains bounded on one side, which require estimates not to give any weight to negative numbers. For estimates, which keep non-negativity into account, see Silverman (1986), pp. 29-32.

---

[61] The kernel estimator is a sum of "bumps" placed at the observations. The kernel function K determines the shape of the bumps, while the window width h determines their width.

[62] The gaussian kernel is based on the standard normal distribution: $K(t) = (1/\sqrt{2\pi})e^{-(1/2)t^2}$

Multivariate density estimation

Consider now to have a sample $X_1,\ldots,X_N$ of independent, identically distributed observations in R^d from a continuous multivariate distribution with probability density function *f*. The multivariate kernel density estimator with kernel *K* and window width *h* is defined by

$$\hat{f}(x) = \frac{1}{Nh^d}\sum_{i=1}^{N} K\left\{\frac{1}{h}(\mathbf{x} - X_i)\right\}$$

The kernel function $K(\mathbf{x})$ is now a function defined for the *d*-dimensional $\mathbf{x}$, satisfying

$$\int_{R^d} K(\mathbf{x})d\mathbf{x} = 1.$$

A kernel I adopt in the estimation is the multivariate Epanechnikov kernel

$$K_e(\mathbf{x}) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-\mathbf{x}'\mathbf{x}) & \text{if } \mathbf{x}'\mathbf{x} < 1 \\ 0 & \text{otherwise} \end{cases}$$

where c_d is the volume of the unit *d*-dimensional sphere: $c_1=2$, $c_2=\pi$, $c_3=4\pi/3$, etc.

The optimal smoothing parameter is

$$h_{opt} = A(K)N^{-1/(d+4)},$$

where $A(K) = 2.40$ for an Epanechnikov kernel with $d = 2$.

Estimation of transition probability matrices

Given a stationary, ergodic Markov chain on the state-space $S = \{1, 2, .., n\}$ with transition probabilities[63]

$$p_{ij} = \Pr\left(X_t \in j \middle| X_{t-1} \in i\right) \quad i, j \in S$$

the initial probabilities

$$p_j^{(0)} = \Pr(X_0 \in j) \qquad j \in S$$

and the stationary limiting distribution $\{\pi_j\}, \pi_j > 0, \sum_{j=1}^{n} \pi_j = 1$, where π_j satisfies

$$\pi_j = \sum_{k=1}^{n} \pi_k p_{kj} \quad \text{and} \quad \lim_{t \to \infty} p_{ij}^{(t)} = \pi_j$$

Let now $\{x(t)$: t = 0, 1, 2, ... , T$\}$ be a realisation of length (T+1). The likelihood function based on this sample is then given by

$$L = p_{x_0}^{(0)} \prod_{t=1}^{T} p_{x_{t-1}x_t} = p_{x_0}^{(0)} \prod_{i,j=1}^{n} p_{ij}^{n_{ij}}$$

where n_{ij} is the observed frequency of the (one-step) transitions from state i to state j in the sample. The set of n^2 transition frequencies $((n_{ij}))$ forms a

[63] These are the probabilities of having a realisation of the variable of interest in state *j*, after a specified transition period, conditional upon prior realisation in state *i*.

sufficient statistic for the transition matrix under estimation.

We now need to maximise

$$\log L = \log p_{x_0}^{(0)} + \sum_{i,j} n_{ij} \log p_{ij}$$

with respect to p_{ij} and subject to the restriction $\sum_j p_{ij} = 1$. Ignoring any information about transition probabilities which may be contained in the initial probability distribution, this yields the Maximum Likelihood Estimator:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i}$$

where $n_i = \sum_j n_{ij}$. This estimator can be shown to be consistent and to have an asymptotic normal distribution[64].

[64] See Basawa and Prakasa Rao (1980).

Figure 1.7 Estimated stochastic kernels

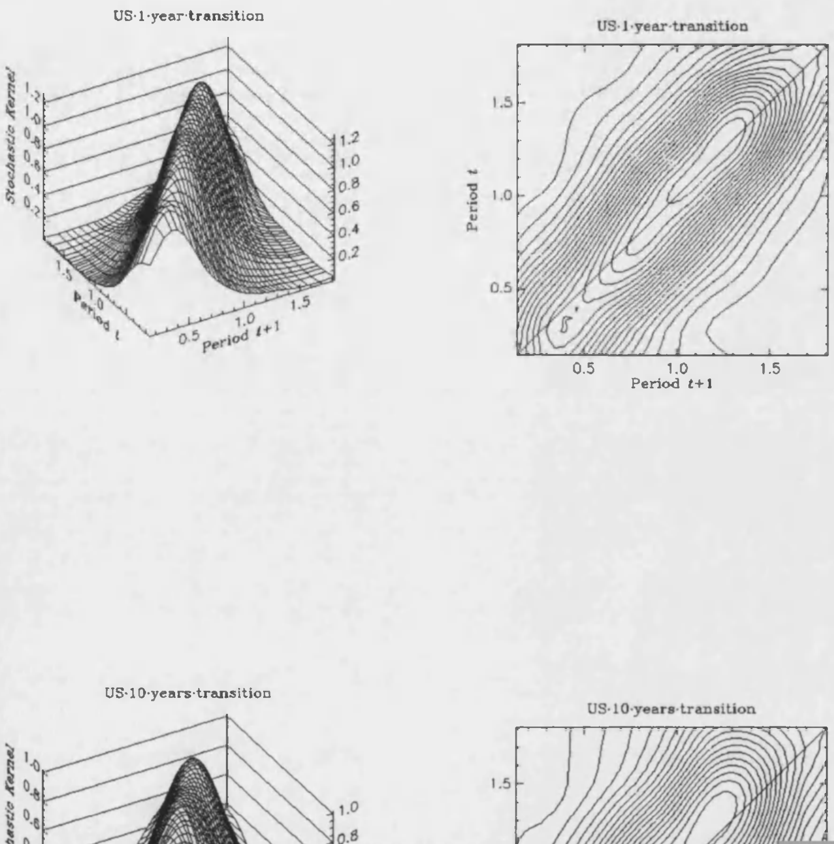


Figure 1.7 (cont.)

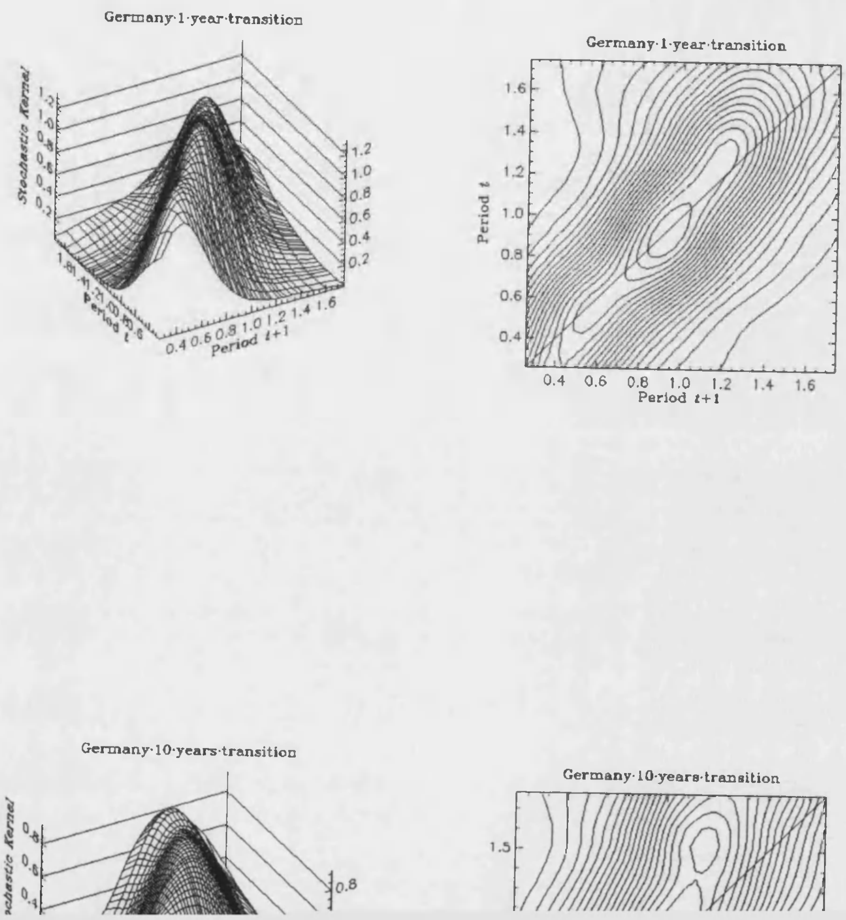


Figure 1.7 (cont.)

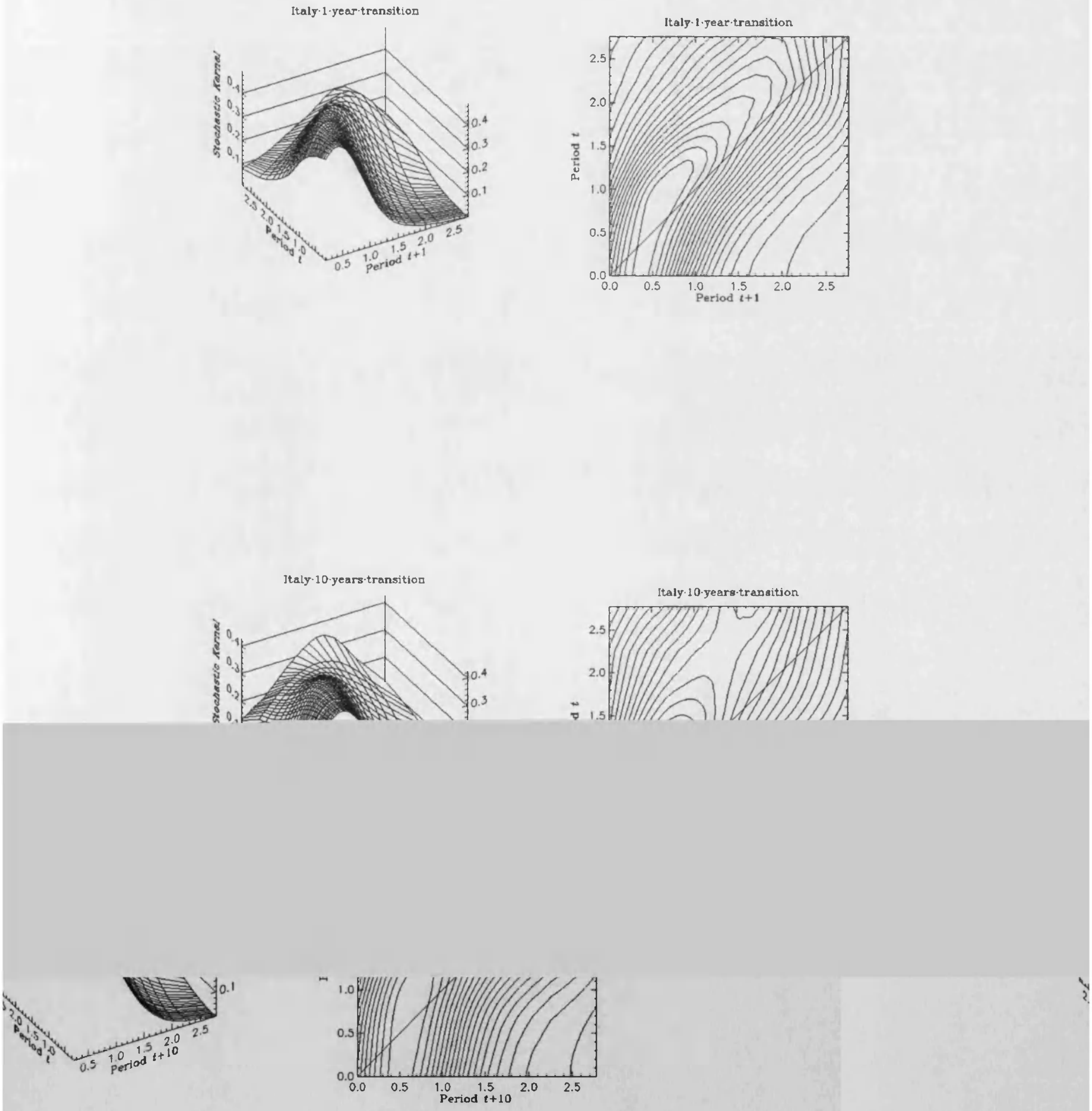


Figure 1.7 (cont.)

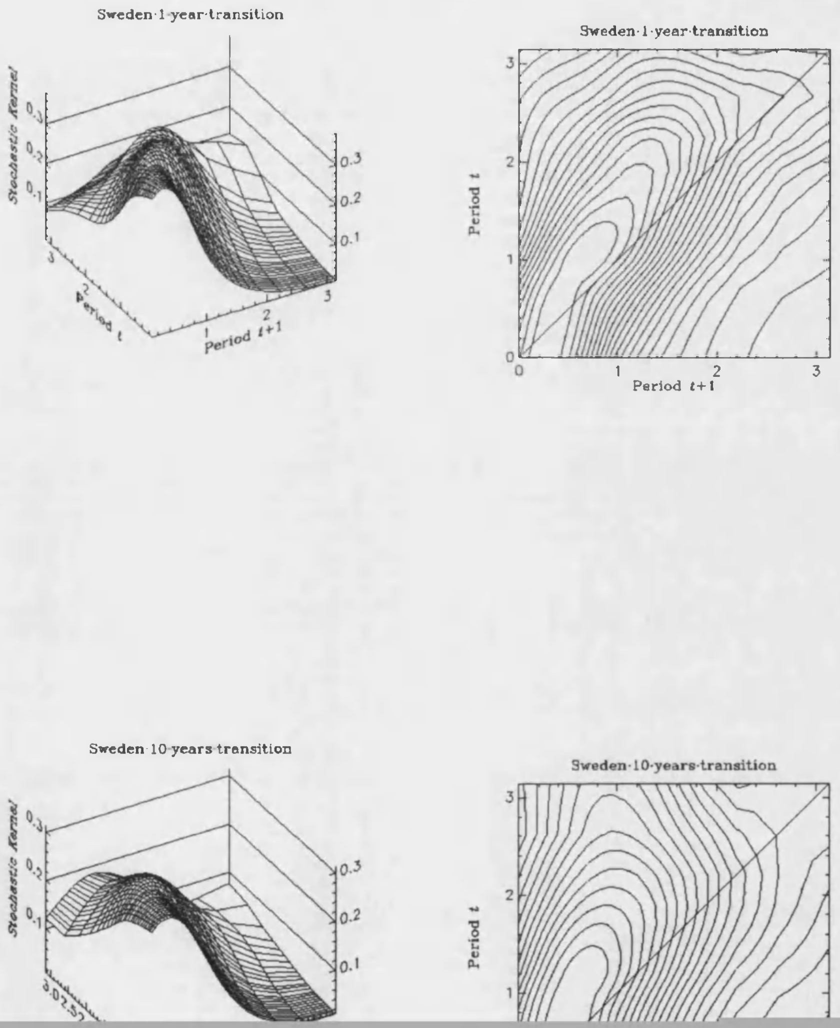


Table 1.3 Five states transition probability matrices

One year transition

US	Upper endpoint				
Number	0.59	0.87	1.11	1.38	118
332	0.65	0.24	0.06	0.02	0.03
344	0.21	0.44	0.22	0.09	0.04
320	0.06	0.25	0.41	0.21	0.08
333	0.04	0.08	0.20	0.48	0.20
323	0.04	0.03	0.08	0.22	0.63
Ergodic	0.199	0.212	0.193	0.206	0.190

Germany	Upper endpoint				
Number	0.62	0.88	1.085	1.32	118
330	0.66	0.22	0.04	0.03	0.04
338	0.22	0.39	0.23	0.10	0.06
322	0.06	0.24	0.38	0.21	0.11
341	0.03	0.09	0.20	0.45	0.24
321	0.05	0.06	0.10	0.25	0.54
Ergodic	0.206	0.199	0.186	0.210	0.199

France	Upper endpoint				
Number	0.51	0.75	1.02	1.38	118
324	0.54	0.20	0.08	0.09	0.08
333	0.20	0.37	0.24	0.16	0.04
335	0.07	0.25	0.38	0.22	0.07
337	0.11	0.13	0.23	0.32	0.22
323	0.09	0.05	0.09	0.22	0.55
Ergodic	0.203	0.202	0.206	0.202	0.188

UK	Upper endpoint				
Number	0.52	0.8	1.07	1.4	118
332	0.48	0.22	0.10	0.09	0.11
329	0.20	0.35	0.24	0.13	0.07
342	0.11	0.23	0.32	0.23	0.11
320	0.10	0.13	0.24	0.30	0.22
329	0.11	0.07	0.10	0.25	0.47
Ergodic	0.202	0.204	0.201	0.200	0.194

Japan	Upper endpoint				
Number	0.34	0.71	1.06	1.52	118
335	0.63	0.24	0.07	0.04	0.01
335	0.21	0.51	0.21	0.06	0.02
333	0.08	0.21	0.45	0.21	0.05
319	0.03	0.05	0.22	0.51	0.19
330	0.02	0.02	0.06	0.18	0.72
Ergodic	0.185	0.207	0.206	0.203	0.200

Ten years transition

US	Upper endpoint				
Number	0.59	0.87	1.11	1.38	118
123	0.53	0.22	0.08	0.11	0.06
115	0.23	0.39	0.20	0.10	0.08
118	0.10	0.29	0.36	0.19	0.06
112	0.05	0.12	0.23	0.43	0.17
122	0.04	0.06	0.11	0.32	0.47

Germany	Upper endpoint				
Number	0.62	0.88	1.085	1.32	118
104	0.56	0.17	0.12	0.04	0.11
141	0.27	0.30	0.20	0.15	0.08
135	0.09	0.18	0.30	0.27	0.16
106	0.04	0.15	0.19	0.31	0.31
104	0.12	0.05	0.15	0.24	0.43

France	Upper endpoint				
Number	0.51	0.75	1.02	1.38	118
115	0.55	0.21	0.06	0.08	0.10
114	0.20	0.29	0.27	0.16	0.08
111	0.10	0.25	0.32	0.22	0.11
130	0.11	0.18	0.26	0.22	0.22
120	0.12	0.09	0.13	0.25	0.40

UK	Upper endpoint				
Number	0.52	0.8	1.07	1.4	118
124	0.40	0.20	0.06	0.19	0.14
111	0.15	0.34	0.19	0.21	0.11
131	0.17	0.21	0.21	0.27	0.15
92	0.18	0.18	0.17	0.27	0.18
132	0.09	0.13	0.17	0.27	0.33

Japan	Upper endpoint				
Number	0.34	0.71	1.06	1.52	118
141	0.52	0.30	0.11	0.04	0.04
99	0.19	0.35	0.25	0.17	0.03
116	0.09	0.24	0.33	0.26	0.08
104	0.03	0.09	0.18	0.46	0.24
130	0.03	0.05	0.15	0.26	0.51

Table 1.3 (cont.)

One year transition (cont.)

Italy	*Upper endpoint*				
Number	0.27	0.56	0.89	1.49	118
338	0.47	0.22	0.12	0.09	0.09
330	0.21	0.42	0.25	0.08	0.03
328	0.12	0.21	0.34	0.23	0.09
320	0.08	0.13	0.23	0.38	0.18
336	0.09	0.02	0.07	0.23	0.59
Ergodic	0.187	0.201	0.207	0.206	0.199

Switzerland	*Upper endpoint*				
Number	0.24	0.55	0.85	1.47	118
332	0.51	0.20	0.11	0.09	0.08
333	0.23	0.45	0.20	0.06	0.05
329	0.12	0.21	0.38	0.21	0.08
329	0.07	0.08	0.24	0.41	0.19
329	0.07	0.05	0.07	0.22	0.59
Ergodic	0.203	0.203	0.202	0.196	0.196

Netherlands	*Upper endpoint*				
Number	0.15	0.57	0.95	1.53	118
335	0.56	0.19	0.09	0.09	0.07
328	0.19	0.46	0.23	0.09	0.03
340	0.09	0.24	0.38	0.21	0.08
324	0.10	0.07	0.22	0.40	0.21
325	0.07	0.03	0.11	0.18	0.62
Ergodic	0.198	0.197	0.205	0.196	0.205

Sweden	*Upper endpoint*				
Number	0.03	0.45	0.9	1.55	118
326	0.61	0.08	0.09	0.11	0.11
335	0.08	0.61	0.21	0.06	0.03
335	0.07	0.21	0.37	0.25	0.10
320	0.11	0.09	0.24	0.32	0.24
336	0.12	0.03	0.11	0.24	0.50
Ergodic	0.199	0.204	0.205	0.195	0.197

Austria	*Upper endpoint*				
Number	0	0.26	0.69	1.57	118
517	0.66	0.04	0.08	0.11	0.12
145	0.18	0.31	0.42	0.08	0.01
330	0.11	0.22	0.41	0.22	0.05
330	0.15	0.02	0.27	0.38	0.18
330	0.19	0.00	0.04	0.20	0.57
Ergodic	0.307	0.089	0.208	0.202	0.194

Ten years transition (cont.)

Italy	*Upper endpoint*				
Number	0.27	0.56	0.89	1.49	118
154	0.26	0.30	0.17	0.18	0.09
107	0.16	0.36	0.29	0.17	0.02
107	0.10	0.21	0.29	0.31	0.09
98	0.09	0.11	0.29	0.26	0.26
124	0.07	0.06	0.14	0.26	0.47

Switzerland	*Upper endpoint*				
Number	0.24	0.55	0.85	1.47	118
118	0.52	0.14	0.09	0.17	0.08
115	0.17	0.41	0.23	0.11	0.09
120	0.17	0.27	0.27	0.20	0.10
110	0.09	0.13	0.26	0.26	0.25
127	0.10	0.07	0.14	0.29	0.39

Netherlands	*Upper endpoint*				
Number	0.15	0.57	0.95	1.53	118
136	0.53	0.18	0.10	0.09	0.10
115	0.19	0.43	0.23	0.11	0.04
123	0.06	0.17	0.31	0.28	0.18
107	0.09	0.12	0.18	0.28	0.33
109	0.08	0.08	0.10	0.21	0.52

Sweden	*Upper endpoint*				
Number	0.03	0.45	0.9	1.55	118
120	0.58	0.10	0.10	0.10	0.12
118	0.04	0.50	0.24	0.18	0.04
111	0.11	0.22	0.35	0.23	0.10
113	0.14	0.11	0.25	0.27	0.24
128	0.16	0.05	0.14	0.30	0.36

Austria	*Upper endpoint*				
Number	0	0.26	0.69	1.57	118
213	0.55	0.04	0.09	0.16	0.16
47	0.13	0.32	0.49	0.06	0.00
101	0.14	0.18	0.41	0.22	0.06
105	0.21	0.07	0.24	0.30	0.18
124	0.20	0.00	0.11	0.31	0.38

CHAPTER 2

GEOGRAPHICAL CONCENTRATION AND SPECIALIZATION DYNAMICS IN DIFFERENT TECHNOLOGICAL FIELDS

2.1 Introduction

The previous chapter presented evidence of substantial mobility in patterns of technological specialization, a finding that contrasts with the hypothesis of technological accumulation and of self-reinforcing mechanisms leading to polarisation in specialization patterns. The empirical results are obtained for a substantial group of OECD countries from the analysis of the evolution of their entire specialization pattern, without taking into account differences across technological fields.

However, knowledge characteristics, cost structures and externality effects may differ across technological fields and this may determine different technology dynamics. A country may experience higher (lower) difficulties in gaining or maintaining a relatively advantaged position in the production of new knowledge (and products in which it is embodied), according to the type of knowledge involved, the share of fixed costs, the importance of learning and coordination effects. In particular, the more the knowledge base in a technology field is complex, cumulative and specific, the more one should expect a country in a relatively advantaged position to be able to reinforce it in the future and one with a relatively disadvantaged position to find it difficult to catch-up.

This chapter looks at the empirical dynamics of specialization in a group of technological fields, in order to find evidence of differences across them, should they exist. First, it identifies the characteristics of the cross-country distribution of innovative activities and successes in a specific technological field. It is well known that the overall cross-country distribution of innovation activities is highly skewed: there are many countries not performing any activity of this type or doing it at a very low scale, and then a relatively small number of countries spending significant amounts of resources for activities aimed at promoting technological change. This is easily seen by looking at patent data, a rich indicator of the output of research and development activities: about seventy percent of the patent applications filed at the European Patent Office comes from innovators resident in either the US, Japan, or Germany.

The first question to be asked then is whether the cross-country distributions of different technology fields have all the same characteristics of the overall one, or if any of them is characterised by a distinctively different pattern. This issue is here studied by looking first at the degree of geographical concentration of each field and of all the technologies included in the field. I then move to the estimation of the cross-country distribution of relative advantages in each field, and observe its evolution in time. This will allow understanding whether technological fields have been characterised by overall stability or by tendencies towards equalisation or polarisation of countries' relative advantages in the production of innovative output.

The second objective of this chapter is then to try to understand whether the evolution of a country's relative advantage in a specific technological field is characterised by stability or if tendencies towards increasing (decreasing) specialization prevail over some range of values. The analysis of intra-distribution dynamics will allow answering questions analogous to the ones asked in the previous chapter, but now related to a specific field of technology. Do countries specialized (de-specialized) in a particular technology field at period *t* show the tendency to remain so in the long run? Is there any field of technology where high de-specialization implies lock-in and, contrary to the generalised findings of chapter 1, high specialization induces positive reinforcing effects because of increasing returns in the creation of new knowledge, thus generating polarisation towards the ends of specialization? Are there elements pointing to the existence of technology specificities that might affect the evolution of countries' ability to innovate persistently in a particular field or to catch-up with actual leaders?

The empirical analysis of this chapter is performed on the following technological fields: electronics, instruments, chemicals and pharmaceuticals, processes, machinery, consumer goods and civil engineering. The fields are quite widely defined; this is due to both data and methodological constraints. It should then be recognised that each field might comprise quite substantial heterogeneity in terms of knowledge base characteristics, learning and externality effects, etc. As a consequence, the analysis reported in the following pages is meant to be exploratory and to complement the empirical analysis of chapter 1.

The chapter is organized as follows. Section 2 provides an overview of the related literature. Section 3 analyses the trends and the degree of geographical concentration characterizing innovation activities in the six broad technological fields. In Section 4 the distribution dynamics approach is again employed to study countries' specialization within specific technological fields and its dynamics in time. Conclusions are drawn in Section 5.

2.2 Sources of increasing returns and differences across fields

Chapter 1 explained how path-dependence in technological change may arise as a consequence of the existence of increasing returns, originating from localised knowledge externalities, which would then induce persistence and polarisation in patterns of technological specialization. However, not all technologies are prone to increasing returns.

Arthur (1994, p. 112) argues that four features of a technology and its social context generate increasing returns:

(1) *Large set-up or fixed costs.* These create a high pay-off for further investments in a given technology. With large production runs, fixed costs can be spread over more output, which will lead to lower unit costs. When set-up or fixed costs are high, individuals and organisations have a strong incentive to identify and stick with a single option.

(2) *Learning effects.* Knowledge gained in the operation of complex systems also leads to higher returns from continuing use. With repetition, individuals learn how to use products more effectively, and their experiences are likely to spur further innovations in the product or in related activities.

(3) *Coordination effects.* These occur when the benefits an individual receives from a particular activity increase as others adopt the same option. If technologies embody positive network externalities, a given technology will become more attractive as more people use it. Coordination effects are especially significant when a technology has to be compatible with a linked infrastructure (e.g., software with hardware, automobiles with an infrastructure of roads, repair facilities and fuelling stations). Increased use of a technology encourages investments in the linked infrastructure, which in turn makes the technology more attractive.

(4) *Expectations.* If options that fail to win broad acceptance will have drawbacks later on, individuals may feel a need to pick the right horse.

Although the dynamics here is related to coordination effects, it derives from the self-fulfilling character of expectations. Projections about future aggregate use patterns lead individuals to adapt their actions in ways that help to make those expectations come true.

Economic theory on innovation and technological change has focused on different factors that are likely to be important in explaining the distribution of technological activities across countries and the evolution of countries' international specialization in technology fields. These factors are related to some of the above mentioned elements and, in particular, to the learning effects. One line of research, already mentioned in the previous chapter, has focused on the nature of the innovation process (Nelson and Winter, 1982; Dosi, 1988; Malerba, 1992). Technological change is viewed as a process leading from technological opportunities to actual innovative effort and then to changes in the structure and performance of industries. The crucial hypothesis is that such process involves the solution to problems and, as such, it is cumulative and shaped by sector-specific technological paradigms that are common across countries.

Technological change is a cumulative process because solutions to technological problems involve the use of information, formal knowledge, and the inventor's specific and uncodified capabilities drawn from previous experience. The search for such solutions is constrained by the technological paradigm because this defines the technological opportunities for further innovations and the basic procedures on how to exploit them, thus determining the directions of innovative efforts (Dosi, 1988). Prior knowledge in a technology field gives firms and countries the ability to understand new information, recognize its value, assimilate it and apply it to commercial ends: abilities which have been collectively named "absorptive capacity" (Cohen and Levinthal, 1990). This has important implications for the innovative performance of firms (and countries) as the empirical results from the previous chapter might also suggest. If a firm (country) has not invested in absorptive capacity (or if a new technological paradigm emerges, thus making the existing knowledge base obsolete), it may not be able to appreciate new opportunities when they subsequently emerge. This may induce lock-in and influence the ability of countries to catch-up with actual technological leaders.

This line of research has mainly focused on characteristics of the knowledge base to explain differences in patterns of innovative activities and their relationship with industrial structures (Pavitt, 1984; Dosi, 1988; Malerba, Orsenigo, 1995 and 1996). The key idea is that the cross-sector distribution of opportunities and capabilities is not homogenous and that appropriability conditions, economic incentives to innovation and the nature of production activities also differ across industrial structures. Within this framework and following a Schumpeterian perspective, Malerba and Orsenigo (1995 and 1996) have proposed a two-way taxonomy of sectors depending on the conceptualisation of technological change as a process of creative destruction[1] or creative accumulation[2].

In the first conceptualisation (*creative destruction*) technological change is a random process, with homogeneous firms drawing from a pool of technological opportunities available to everybody. The monopoly power generated by a successful innovation is soon eroded by a competitor's success, hence turnover is high and the typical innovative firm is small and often newly established. In the second conceptualisation (*creative accumulation*) technological change is originated by a knowledge base with strong tacit components, which may also be specific to individual firms and applications. Innovation is here the result of in-house accumulation of knowledge and capabilities and comes mainly from established and long experienced firms

[1] From the interpretative model Schumpeter discussed in "The Theory of Economic Development" (1912).
[2] From the interpretative model Schumpeter discussed in "Capitalism, Socialism and Democracy" (1942).

operating at the technological frontier. Disruption of their leadership requires drastic changes in the technological paradigm, which make the knowledge accumulated by the innovative leaders obsolete.

Whichever conceptualisation best describes technological change in a sector, one should expect technological specialization to be associated with specific characteristics of innovative firms, concentration of innovative activities and degree of technological entry. Empirical evidence (Malerba et al. 1997) has found that relative technological advantages of a country tend to be higher in sectors characterised by the existence of a competitive core of large firms which innovate systematically over time, thus signalling the importance of "creative accumulation".

Overall, economic literature on technological change has argued that the evolution of patterns of innovative activities may differ across technological fields because technological opportunities and the importance of being innovative as a source of competitive advantage may themselves vary significantly across fields. As a consequence, international technological specialization of countries may be characterized by technology specificities, that is the maintenance and future development of a country's innovation capabilities in a field may depend upon characteristics of the field itself.

Among the factors generating increasing returns, other authors have stressed the importance of location specific advantages to innovative performance related to the structural characteristics of the economic environment. Both Krugman (1991) and Arthur (1994) point to the role of increasing returns in the spatial location of production. Given the importance of physical proximity in many aspects of economic life, agglomeration effects are widespread. That is, initial centres of economic activity may act like a magnet and influence the location decisions and investments of other economic actors. Established firms attract suppliers, skilled labour, specialized financial and legal services, and appropriate physical infrastructure. The concentration of these factors may in turn make the particular location attractive to other firms that produce similar goods. So do social networks, which facilitate the exchange of information and expertise. In particular, social networks of individuals cutting across companies' boundaries and university campuses have been often held responsible for the circulation of valuable information and for filling the air (of both Marshallian districts and hi-tech clusters) with bright new ideas (for a survey: Breschi and Lissoni, 2001).

2.3 Trends and geographical concentration in technology fields

To evaluate geographical concentration and the extent of *mobility* and *persistence* of international technological specialization in a specific field I employ the same data used in the previous chapter: all patent applications filed at the European Patent Office (EPO) between 1982 and 1996.

Recall that each patent is assigned to a specific technology defined according to the International Patent Classification (IPC), which is hierarchical, primarily concerned with the technological characteristics of the innovation, and designed to represent the whole body of knowledge, which may be regarded as proper to the field of the invention.

Here the IPC classes have been grouped into 30 technologies, according to a classification developed by the Fraunhofer Institute for Systems and Innovation Research (FISIR) and reported in the Appendix (Table 2.4). These can in turn be grouped into six broad technological fields: *electronics* (technologies 1 to 5), *instruments* (technologies 6 to 8), *chemicals-pharmaceuticals* (technologies 9 to 14), *processes* (technologies 15 to 20), *machinery* (technologies 21 to 28), and the residual field including the two remaining technologies (consumer goods, 29, and civil engineering, 30).

The world shares of patents at the EPO in the technological fields defined above are characterised by different patterns. The share of electronics grows from 19.3 percent in 1982 to almost 26 percent in 1996, with higher growth rates after 1988. Also the share of instruments increases from 13.8 percent to almost 16 percent between 1982 and 1993, but then it starts declining. On the contrary, the world percentage share of chemicals and pharmaceuticals declines monotonically from 23 percent in 1982 to around 19.5 percent in 1996. Process technologies' share remains quite constant throughout the period, with an average share of 14.8 percent. Finally, the percentage share of the machinery field decreases from 21 to 19 percent throughout the sample period. Overall, the average yearly growth rate between 1982 and 1996 is significantly positive only for electronics (+2.2 percent) and significantly negative for processes and chemicals-pharmaceuticals (-1.7 and -1.3, respectively)[3].

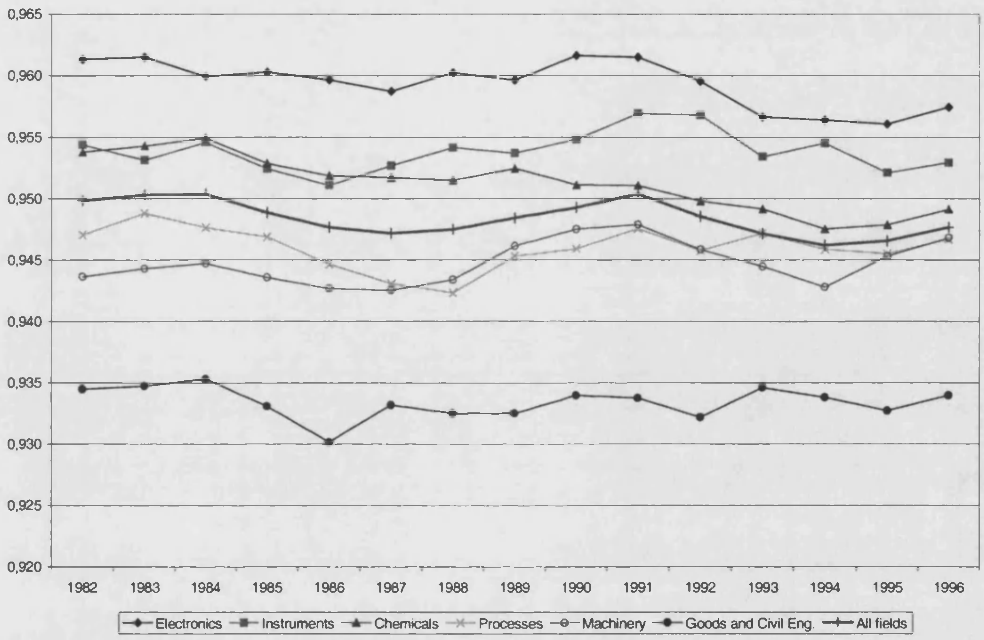
Geographical concentration in the production of innovative output in each technology field can be measured using, once again, the Gini coefficient. Recall that it ranges between zero and one and it is equal to twice the area below the Lorenz curve, here obtained by plotting country i's share of patent applications in field j (i.e. $P_{ij} / \sum_i P_{ij}$, where P_{ij} is the number of patent applications in field j from country i) against the cumulative percentage of countries (ordered by increasing share of applications in the field). This implies comparing the distribution of shares to a uniform distribution: the Gini coefficient gives then a measure of absolute specialization or geographical concentration.

Figure 2.1 shows the evolution of the standard Gini coefficient for each of

[3] The increase in the electronics percentage share of all patent applications at the EPO is mainly due to the significant growth of telecommunications (technology 3) and information technology (technology 4), even if the older and mature electrical engineering field still has the largest share within the group. The drop in the chemical and pharmaceutical technologies' share is mainly caused by the constant and sustained decline in organic chemistry (technology 9), which falls from a 10 percent share in 1982 to only 5 percent in 1996. On the contrary, the shares of pharmaceuticals (technology 11) and of biotechnology (technology 12) have been constantly increasing throughout the period.

the six technology fields and for the distribution of countries' shares of all the patent applications (i.e. applications in all fields) at the EPO. The extremely high level of the index in all cases confirms that the production of innovations is confined to a few geographical areas, regardless of the type of technology.

Figure 2.1 The Gini coefficient calculated for technological fields



Overall, Figure 2.1 shows that geographical concentration is fairly stable similar in the last decade of the sample[4].

Table 2.1 The average standard Gini coefficient (average degree of geographical concentration) and the average growth rate over the period 1982-1996 for the 30 FISIR technologies.

		Average growth (1982-1996)	
		Positive	*Negative*
Standard Gini (Degree of geographical concentration)	*Above* average	2, 3, 4, 5, 6, 7, 8, 11, 12, 22, 26	1, 9, 10, 15, 17, 27
	Below average	20, 24, 29, 30	13, 14, 16, 18, 19, 21, 23, 25, 28

Note. "Standard Gini" refers to the average standard Gini coefficient in the period 1982-1996, calculated for each technology. For each technology, it is "above average" if its value is above the period average value of the standard Gini coefficient calculated from overall countries' patenting shares (i.e. their total patenting)

Once the Gini coefficient is calculated for each of the FISIR technologies it is possible to distinguish between technologies whose share has been growing in time versus declining technologies. Table 2.1 synthesizes for the 30 FISIR technologies the information on the average standard Gini coefficient (average degree of geographical concentration) and the average growth rate over the period 1980-1996. All the electronics technologies have been growing, with the only exception of technology 1 (Electronic devices and Electrical engineering), that remained stable throughout the period[5]. For all of them, the average Gini coefficient is above the overall average[6], thus showing a higher than average

[4] When the Gini coefficient is calculated for each of the FISIR technologies, it shows that, within the electronics field, innovation activities in information technology (technology 4) and semiconductors (technology 5) appear as the most geographically localised, while those in the electrical engineering and telecommunications technologies are more dispersed. In the instruments field, it is the optics technology the one with the highest degree of concentration (technology 6). In chemicals, macromolecular chemistry and polymers technologies (technology 10) are the mostly concentrated, while, not surprisingly, technologies related to food and agriculture (technology 14) are the most diffused.

[5] Its average growth rate is negative, but negligible: -0,000003.

[6] This is the time average of the standard Gini coefficient calculated over patent applications in all fields.

degree of geographical concentration of innovative activities in this field. Also instruments include fairly concentrated technologies, two of which (out of the three belonging to this field) have been growing significantly (Optics and Medical engineering, i.e. technologies 6 and 8).

Technologies belonging to the chemicals field mostly show a relatively high level of geographical concentration; among these, pharmaceuticals and biotechnology (technologies 11 and 12) are characterised by a positive growth rate. The machinery field includes technologies with different characteristics, while almost all process technologies have negative average growth rates[7] and a relatively low degree of geographical concentration.

Overall, the analysis of this section has shown that innovation activities are extremely localised in any technological field, and are particularly so in some fast growing, R&D intense technologies (as IT, optics, pharmaceuticals, biotechnology, etc.).

2.4 Distribution dynamics

In order to provide evidence on the dynamics of a country's specialization in each field I use the *RTA* index introduced in chapter 1. The distribution of the degree of international technological specialization in a field at any one point in time can be characterised by the distribution of *RTA* across countries[8].

[7] The only exception is technology 20 (Environmental technology) which has been slightly growing, probably as a consequence of the growing attention to environmental issues and pollution problems.

[8] With reference to the discussion of the *RTA* vs. the *TRCA* index presented in chapter 1, note that this gives a further reason for employing the *RTA* rather than the *TRCA* index. The latter suffers from the problem of across-time ranking, i.e. the mean of the index varies both across time and countries: the same value or change of the *TRCA* index, interpreted as the relative importance for a field in a country, can have different meaning for different countries and for the same country at different time periods. By stabilizing the mean to one, the Proudman and Redding normalisation makes the index comparable across countries and across time. For an extended discussion on the Balassa and Proudman and Redding indexes see also De Benedictis

Hence, evaluating the dynamics of international technological specialization in a field over time requires an analysis of the evolution of the entire cross-section distribution of *RTA*. On one hand, it is interesting to ask whether technological fields have been characterised by overall stability or tendencies towards equalisation or polarisation of countries' relative advantages in the production of innovative output. This may be evaluated by analysing the evolution of the *external shape* of the cross-country *RTA* distribution.

On the other hand, it is also relevant to understand whether the evolution of an individual country's relative advantage in a specific technological field is characterised by relative stability or if tendencies towards increasing (decreasing) specialization prevail over some range of values. For example, it is interesting to know if very high de-specialization implies lock-in and if relatively high specialization induces positive reinforcing effects because of increasing returns in the creation of new knowledge, thus generating a concentration of probability mass at the extreme ends of specialization. Studying the *intra-distribution dynamics* for each technology field allows addressing these questions and then performing comparisons across fields.

The evolution of the *RTA* distribution over time is here modelled adopting the distribution dynamics approach, presented in the previous chapter. The analysis reported in the following pages will address the first issue above (*changes in the external shape of the distribution*) by estimating the distribution of *RTA* across countries for each technology field. It will then move to the *persistence vs. mobility* issue, where persistence will again be interpreted as a measure of the probability of remaining in the state in which a country initially is. Namely, if a country is specialized in a field, the aim here is to know what is the probability that it remains specialized as time goes by.

Chapter 1 employed the same techniques to derive the empirical dynamics

---

and Tamberi (2001).

of technological specialization patterns in industrial countries. There patents are taken as an indicator of countries' sectoral distribution of research output: their distribution across technologies summarised the technological frontier of a country and its pattern of technological specialization at a specific point in time. However, intra-distribution mobility is studied with no reference whatsoever to differences across technological areas or fields. The techniques are devised to follow the path of a particular observation (i.e. the value of the *RTA* index for a country in a technology), considering it as the outcome of one unique process. In other words, that chapter implicitly assumed that the underlying stochastic process is the same in all the technologies for a given country. The aim of this chapter is precisely to find evidence in favour or against that assumption and to check for possible technology specificities in the dynamics of country specialization[9].

A final remark is due about the high aggregation level adopted. This is the unfortunate consequence of the trade-off between the informative content of the techniques employed and the constraints they impose. They require a relatively high number of non-zero cross-section observations, which cannot be obtained for low aggregation levels, precisely because the distribution of patent applications (and innovation activities) across countries is highly skewed and it is even more so the more restrictive is the definition of technology field adopted. Nevertheless, even at a high aggregation level the employment of these innovative techniques can provide a valuable preliminary insight into the dynamics of technological specialization in different technology fields.

All the probability density functions, Markov stochastic kernels and transition probability matrices, presented in this chapter, have been estimated from EPO data using Danny Quah's TSRF.

[9] Note, however, that I now have to assume that the underlying stochastic process is the same in all the countries for a given field. Unfortunately, this is an assumption I cannot relax.

2.4.1 Changes in the shape of the cross-sectional distributions

The first step is to estimate the distributions across countries of the specialization index (*RTA*) in each of the six broad technological fields[10]. By drawing these distributions it is possible to see whether any of them is already characterised by polarisation towards the extremes, with a possibly large number of de-specialized countries and a few highly specialized ones.

The distributions are obtained by pooling the observations of three five-year sub-periods: 1982-1986, 1987-1991, and 1992-1996. There are 126 countries applying for patents at the EPO over the sample period; however, here the data have been "censored" in the following way: for each technology field all the countries which never apply for a patent in the field have been excluded. This means that a different number of observations have been used to obtain the cross sectional distribution of the six fields, corresponding to the number of countries patenting at least once in the field over the period 1982-1996. These are 76 for Electronics, 81 for Instruments and for the Consumer Goods and Civil Engineering field, 95 for Chemicals & Pharmaceuticals, 94 for Processes and 93 for Machinery.

The estimated cross-sectional distributions for the 1982-1986 sub-period are strongly influenced by the still low patenting activity at the EPO. Indeed, for all technology fields the cross-country distribution is characterised by a very high peak at zero and a small residual density over the interval (0,6]. Therefore, it is only interesting to look at the following two sub-periods (see Figure 2.2 in the Appendix). Since they are so close in time, one should not expect to find any striking difference in the shape of the probability distribution function between the two. Rather, it would be interesting if the estimation pointed out some systematic difference among the technology fields or groups of them.

---

[10] All the densities are estimated by Gaussian kernel smoothing, taking non-negativity into account and following the procedure and automatic bandwidth choice from Silverman (1986: 2.10 and 3.4.2).

Figure 2.2 shows, the cross-sectional distributions of all the fields have a very high peak at zero: even if the countries which never apply for a patent in the field have been excluded, each year there is still a large number of countries not innovating in the field. Each year, in any of the six broad technology fields around 50 percent of the countries are not applying for a patent, the percentage being higher for electronics and instruments. This confirms that innovation activities are not widespread, but are instead fairly localised, a fact which is not surprising, since those activities require a considerable amount of resources to be spent on R&D, and supply and demand conditions which are not present everywhere in the world[11].

All the technology fields analysed, but electronics show then a tendency for a second peak to appear. This is very clear for instruments, chemicals and pharmaceuticals, processes and machinery. The densities of these four fields are characterised by a distinctive peak around the mean value of the specialization index (slightly below for instruments). This implies that each year, together with a large number of countries not patenting at all in the field, a good number of countries have no marked specialization or de-specialization in the same field. Note how the second peak is wider in the distribution of machinery, thus signalling a greater dispersion around the value of 1. Finally, in consumer goods and civil engineering, the second peak is much lower (especially in the second period) and localised between 1 and 1.5. This agrees with the picture of this field given in the previous section: there is a quite large number of (small) countries specialized in these fairly standardized and mature technologies.

Electronics has distribution characteristics different from the other fields. The peak at zero is very high and remains so even in the second period, while it

[11] Recall that the main point in the "technology-push vs. demand-pull" debate is that invention is a response to both technology and profit opportunities. These, in turn, depend on the characteristics of the underlying knowledge base and on the infrastructures promoting its diffusion, on the industry structure, on the size, needs and elasticity of the market.

becomes much lower in all the other fields. There is then no clear peak around the value of 1; on the contrary, a lot of the probability mass remains below that level. This implies that the electronics field is characterised by the presence of countries with a clear specialization or de-specialization in electronic technologies, with a marked tendency for the de-specialized countries to outnumber the specialized ones[12].

Indeed, among the specialized countries there are only two of the G5 members, Japan and the US, and then the Netherlands and Finland (from the beginning of the nineties)[13], and a series of "small" (in terms of their patenting activity), but persistently specialized countries, such as Korea, Taiwan, Singapore. The value of the specialization index of these countries can be quite dispersed (over the range above 1) and generates the pattern characterising the cross-country distribution described above.

2.4.2 Intra-distribution dynamics

To evaluate the extent of mobility in countries' specialization within technology fields, one-year and ten-year transition probability matrices have been derived for each field. These are 3x3 matrices describing the transitions between three different states of the process guiding the evolution of the specialization index in each field. State 1 (*zero-patent state*) is characterised by no applications for patents in the field and is defined by a single point corresponding to a value of zero of the specialization index. A value of the *RTA* index below or equal to 1 defines state 2 (*de-specialization state*), while

[12] The average year standard deviation of the specialization index in electronics is 0.79, which is below the standard deviation of the index in the other fields. This is 0.89 for Instruments, 1.29 for Chemicals, 1.22 for Processes, 1.19 for Machinery and 1.22 for the Consumer Goods and Civil Engineering field.

[13] Recall these are the home countries of Philips and Nokia, respectively. This may be of primary importance in determining their high specialization level.

state 3 (*specialization state*) includes all the observations on values of the specialization index above 1. State 2 and 3 correspond, respectively, to a country being de-specialized and specialized in the technology field.

Table 2.3 (in the Appendix) reports the estimation results for the so constructed transition probability matrices. Recall the interpretation of each panel in the table. The numbers in the first column are the total numbers of country-year observations beginning in a particular cell, while the first row of numbers denotes the upper point of the corresponding grid cell. Row j includes the estimated probability of remaining in state j (i.e. the element (j,j) of the matrix) and of moving from state j to state s (i.e. the element (j,s) of the matrix). Each row is therefore the probability distribution of the RTA index in a particular field across countries at time t, given its value at time t-1. As such, the row sum is equal to one. The final row of each panel gives the implied ergodic distribution and provides the asymptotic unconditional probability of being in one state, that is the probability of being in one state regardless of the initial state.

The first thing to notice is the high persistence in the zero-patent state over both one and ten-year transition periods. Mobility out of the zero-patent state is indeed low and seems mostly directed towards the specialization state. This result does not have to be overstated since it is probably the consequence of small countries having some exceptionally good result from their research in a technology field.

The lowest persistence is, instead, found in the specialization state, with a tendency of countries originally specialized to de-specialize or move to the zero-patent state in similar proportions. Also this result requires a qualification, since it can arise, at least partially, as a consequence of the just mentioned occasional patenting phenomenon and, possibly, of a generalised decline in the level of specialization due to the increasing number of applications and applicants at the EPO[14]. Persistence in the top state is highest for the consumer goods and civil engineering field, both over one- and ten-year transition periods. The electronics field, instead, is characterised by very high persistence in the de-specialization state and the lowest probability to move from the zero-patent state to the specialization state, again over both transition periods. In other words, in the electronics field it seems to be mostly difficult to move out of a relatively disadvantaged position. Note also that the number of observations starting in state 3 is lower than in the other cases: this confirms, once again, that there are only few countries specialized in the electronics technologies. This feature is also present in the instruments field, while all the other technologies have the opposite pattern, i.e. the number of observation starting in the top state is greater than that of those starting in the de-specialization state.

The ergodic distribution implied by the one-year transition process has always a peak in the zero-patent state: most countries either do not do research or their innovative activities are not successful. There is then a prominence of the top state over the middle one in all the fields but, again, electronics and instruments, thus confirming the tendency to have very few countries specialized in those technology fields.

The evidence presented is affected by the inclusion of countries of vastly different sizes in terms of patenting activity. When the largest of a few of the largest economies increase their share in patenting in a field, then more of the smaller economies must be loosing shares in this industry. Consequently, the Maximum likelihood estimator assigns a larger weight to the more frequently observed losses of the more numerous smaller economies than to the corresponding gains of one or very few big economies. A stationary distribution with the concentration at one end may thus reflect a monotone

[14] The average annual growth rate of patent applications at the EPO is 6.7 percent between 1982 and 1996, with actual growth rates higher than the average before 1990.

trend in the specialization dynamics in the largest economy in the sample (the United States). One way of dealing with the inconvenience caused by the great disparity in the sizes of countries in the sample is to estimate transition matrices for fractile Markov chains, as in the previous chapter. The fractile transition matrix may help on this matter because, instead of using an arbitrary grid to discretise the continuous state space of the specialization indicator, one only needs to fix a set of increasing, non-redundant probabilities, equally spaced in the unit interval, and then let this determine for each period t a corresponding set of quantiles[15]. However, estimated fractiles reveal about the same degree of persistence on the main diagonal as observed in the previously reported non-fractile Markov chain estimates. Further, for all the fields here examined no significant time trend is found in the simple linear regression of the interquartile range on time: there is no evidence of polarisation of the ergodic distribution towards the extremes.

Mobility indices (Shorrocks 1978; Geweke et al. 1986; Quah 1996b) have been calculated from the one-year transition probability matrix of each technological field. Recall that each of these indices attempts to reduce information about mobility from the matrix of stationary one-year transition probabilities M to a single statistic.

Mobility indices are reported in Table 2.2: for the sake of comparison, a transition matrix has also been estimated from transitions in all fields and mobility indices have been calculated for this matrix as well. Table 2.2 shows how electronics is, according to all the indices, the least mobile field[16]. This confirms that a country's specialization in electronics displays stronger tendency to persist around its original level compared to the other fields, thus suggesting that the amount and quality of prior experience in electronics is a

[15] See Appendix B in chapter 1.

[16] It is the only one whose indices take values all below the ones calculated from the transition matrix for all fields.

tighter constraint to future developments.

The finding of high persistence and low mobility of countries' specialization indicators in electronics is consistent with the results in Stolpe (1995). His results are even stronger in that he finds that the Markov chain estimates for the transitions of specialization indicators in electronics[17] even divide into two ergodic sets (Table A9, p. 232). This suggests that countries either belong to a camp with relatively little innovative activity in electronics or to a camp with a lot of such activity, but that they do not cross the boundary between these two camps[18].

Table 2.2 Mobility indices.

	μ_1	μ_2	μ_3	μ_4	M_5
Electronics	0.315	0.178	0.320	0.553	0.187
Instruments	0.440	0.266	0.420	0.738	0.218
Chemicals	0.370	0.293	0.370	0.623	0.223
Processes	0.470	0.353	0.465	0.758	0.267
Machinery	0.420	0.320	0.390	0.683	0.277
G.&C.I.	0.340	0.255	0.420	0.576	0.218
All fields	0.380	0.271	0.380	0.635	0.229

$\mu_1 = \frac{n - tr(M)}{n-1}$; $\mu_2 = \sum_k \pi_k \sum_l p_{kl}|k-l|$; $\mu_3 = \frac{n - \sum_j |\lambda_j|}{n-1}$; $\mu_4 = 1 - |\det(M)|$; $\mu_5 = 1 - |\lambda_2|$

As in the previous chapter, the asymptotic properties of first-order Markov chains derived in Anderson and Goodman (1957) can be used to test for the

[17] Stolpe (1995) used US patent data and assigned patent counts to broad technological fields corresponding to industries of the international industrial classification scheme, ISIC. In his study electronics technology refers to Radio, Television and Communications Equipment (RTVC).

[18] Moreover, in the fixed state Markov chain estimates (Stolpe, 1995, footnote 121 on p. 142) the highest state of the RTVC specialization indicator seems to be absorbing.

statistical significance of the similarities between the dynamics in different technological fields. The test has been implemented to perform pair-wise comparisons between the technological fields in both directions (i.e. using alternatively the transition probability matrix of each of the technological fields in the pair as the null), and has revealed no statistically significant similarity between the one-year or ten-year transition processes of any pair of the six technology fields.

In chapter 1 it was emphasised that the transition probability matrix gives less information than the corresponding stochastic kernel (i.e. the stochastic kernel representing the same Markov process for the same transition period). From a transition probability matrix one may learn that if the process is currently in state k (i.e. if the value of RTA_{ij} belongs to the range identified by state k) then in the next period (however long) it will be in state h with probability p_{kh}. On the contrary, the stochastic kernel gives the entire next period probability distribution of RTA_{ij} over the whole range of possible values, given the exact value it has in the current period. In other words, the transition probability matrix represents the probability of moving from state k to state h with a single number and consequently hides all the transitions taking place within any grid cell[19].

For each technological field, two Markov stochastic kernels have been estimated to represent the conditional probability distributions of the *RTA* index for one-year and ten-year transition periods. For each cross-sectional unit (i.e. each country-field observation) a time series from 1982 to 1996 is available. The stochastic kernels here presented were estimated considering every time series as an independent realisation of the same process and by pooling all the observations on the transitions between period t and period $t+k$

[19] This might be particularly true for the last state (here the 'specialization state'), which is particularly wide.

(where k is equal to 1 and 10)[20].

Figure 2.3 (in the Appendix) shows the stochastic kernels for 1-year and 10-year transitions in the *RTA* data for each of the six technology fields.

The dynamic pattern of electronics appears interestingly different from the others. The one-year transition kernel is mainly centred along the main diagonal, where it is characterised by high density between zero and one, and by a peak at values above one, corresponding to the specialized countries, which remain so in the next period. There is then a peak along the *period t* axis at values of the *RTA* index just above 2. It signals shifts from specialization to no patents at all in the field and refers to the occasional patentees' phenomenon already mentioned (i.e. "small" countries with a one off successful innovation in the field). This last part disappears in the estimated ten-year transition stochastic kernel, which is nevertheless characterised by the same pattern as the one-year transition kernel along the main diagonal[21].

The one-year stochastic kernel for instruments is again characterised by most of the density lying along the main diagonal and some along the *period t* axis. This density along the main diagonal has distinctive peaks at zero and just below 1: countries highly de-specialized or with just below average specialization in the field show the tendency to maintain that position in the following period. As for electronics, the ten-year transition kernel reduces solely to the pattern along the main diagonal. The estimation results for the machinery field are fairly similar, but the density along the *period t* axis is still significant even over the longer period horizon.

The one-year stochastic kernel representing the evolution pattern of the chemicals and pharmaceuticals field is quite similar to the corresponding one of the instruments field, but it also has some density at very high values of the

[20] See chapter 1 for the definition, interpretation, estimation and presentation of Markov stochastic kernels.

[21] The long-run shape of the stochastic kernel is consistent with the results in Stolpe (1995) for the electronics sector.

specialization index, always along the main diagonal. The ten-year stochastic kernel has the same features as the one-year transition one, except this last part. There is instead some density to the right of the main diagonal indicating that some countries already specialized in the chemical field may, with some probability, specialize even more in the long run.

The one-year transition estimation results for the process technologies are very similar to those for chemicals and pharmaceuticals, while the ten-year transition stochastic kernel lies, once again, along the main diagonal and is characterised by a peak at zero and one at 1 (just above the main diagonal). Finally, the consumer goods and civil engineering field is characterised by a quite messy pattern in the one-year transition kernel. However, there is always a good part of the density along the main diagonal, with the usual peaks around zero and one (slightly above it), and some of the density lying along the *period t* axis. As it happens for the other fields, the ten-year transition kernel is much simplified and similar to that estimated for the chemicals technologies.

To sum up, there seems to be a tendency towards persistence in the zero-patent/zero-specialization state and also in a situation of not marked specialization/de-specialization, a tendency clearer over longer periods. By contrast, and with the exception of electronics, there appears to be a tendency of high specialization levels to revert to the mean, a result consistent with those obtained in chapter 1 at the country level. This result is, however, also the consequence of including a large number of small countries, whose specialization pattern is subject to wide fluctuations in time simply because of their limited innovative activities.

2.5 Conclusions

This chapter has first obtained information on the characteristics of the cross-country distribution of innovative activities in six broad technological fields: electronics, instruments, chemicals and pharmaceuticals, processes, machinery, and consumer goods and civil engineering. It has shown that innovation in all the fields is highly geographically localised and that the degree of concentration is fairly stable in time. This is mostly high in the electronics field and, as expected, low in the consumer goods and civil engineering field.

The estimation of the cross-country distribution of the specialization index has then revealed that the distributions of all technology fields, but electronics, are similar in that they all have a clear peak around one, thus signalling that, each year, in each field, a good number of the countries applying for patents does not show a marked specialization/de-specialization, a result that is probably due to the high level of aggregation adopted.

Intra-distribution mobility analysis has shown a strong tendency to persistence in the zero-patent state: a country which today does not have enough experience (i.e. accumulated knowledge) and capabilities in a technology field to be an innovator will find it increasingly difficult to innovate in the future. That this is indeed the case can be seen also from the asymmetry of movements between the zero-patent and the specialization states. A significant part of the mobility in both directions is due to the presence of countries with a very low level of patenting activity, which are very specialized and whose specialization pattern exhibits jumps even over short periods of time. Nevertheless, movements from the top state to the bottom state are more frequent than those in the opposite direction. Once again, this could signal the relevance of the absorptive capacity argument of Cohen and Levinthal (1990).

Persistence is lower, but still present for specialization levels around the mean, while high specialization levels show the tendency to revert towards lower values. This result confirms the asymmetric character of mobility already found in the country analysis of chapter 1[22].

The estimated stochastic kernels show the tendency towards persistence of within field countries' specialization levels around or below the mean more clearly over long transition periods. This is also because the occasional patentees or small countries phenomenon of highly discontinuous patenting activity affects the short run dynamics: over a longer horizon the noise in the process induced by these observations jumping up and down fades away.

There are again elements pointing to the peculiarity of the electronics field with respect to the other technology fields, a result which requires further attention.

Among industrial sectors, the hypothesis of technological accumulation has greater relevance for those classified as high-technology industries, which make intensive use of technological innovation. It has, therefore, greater relevance for those technology fields related to those industries, i.e. presumably those fields where the pace of innovation is relatively fast and to which a large share of R&D resources are devoted.

The empirical analysis of this chapter was specifically aimed at detecting differences across technology fields in the pattern and dynamics of countries' specialization. The ideal setting for such an analysis would have been one at a low level of aggregation, so to be able to compare highly dynamic technologies to old and stagnant ones. Unfortunately, the data requirement for the type of analysis performed forced it towards a much higher level of aggregation, thus limiting its scope. However, this limitation makes the results obtained for the electronics technology even stronger. The estimated distribution of the degree

[22] In the estimated transition probability matrices, the probability of remaining in the specialization state is always below those of remaining in the zero-patent or in the de-specialization states (the only exception is the consumer goods and civil engineering field). Note, however, that the specialization state includes all the values of the index above one, hence it implicitly over-estimates persistence for very high values (e.g. close to 2) and under-estimates persistence for values close to 1. This emerges clearly from the estimated stochastic kernels.

of countries' specialization in electronics is different from those obtained for the other technology fields: electronics has the highest degree of concentration and is characterised by a significant presence of "small" countries among those specialized. Furthermore, it appears as the least mobile technology field according to all the mobility indices calculated.

In a Schumpeterian perspective, the high persistence of the electronics technology may signal the importance of "creative accumulation" for technologies belonging to this field. In such a setting high relative technological advantages result from continuous innovations stemming from a competitive core of large firms (Malerba et al. 1997). This is the case of electronics: of the few countries specialized in this technology, some are characterized by the presence in the field of a few large firms responsible for a large share of electronics patents (the Netherlands, Finland and Japan are the most remarkable examples).

Appendix: Tables and Figures

Table 2.3 Estimated transition probability matrices

ONE YEAR TRANSITION

Electronics	*Upper endpoint*		
Number	0	1	6
1323	0.93	0.03	0.04
393	0.09	0.85	0.07
174	0.26	0.15	0.59
Ergodic	0.665	0.236	0.100

Instruments	*Upper endpoint*		
Number	0	1	6
1277	0.91	0.02	0.06
345	0.07	0.74	0.19
268	0.28	0.25	0.47
Ergodic	0.647	0.203	0.151

Chemicals	*Upper endpoint*		
Number	0	1	6
1170	0.89	0.03	0.08
335	0.07	0.75	0.18
385	0.23	0.15	0.62
Ergodic	0.591	0.190	0.219

Processes	*Upper endpoint*		
Number	0	1	6
1202	0.88	0.03	0.09
267	0.09	0.61	0.30
421	0.24	0.20	0.57
Ergodic	0.596	0.164	0.241

Machinery	*Upper endpoint*		
Number	0	1	6
1207	0.88	0.02	0.09
259	0.12	0.67	0.22
424	0.24	0.15	0.61
Ergodic	0.622	0.148	0.230

G.&C.Eng.	*Upper endpoint*		
Number	0	1	6
1305	0.92	0.01	0.08
97	0.07	0.65	0.28
488	0.19	0.06	0.75
Ergodic	0.667	0.060	0.273

TEN YEARS TRANSITION

Electronics	*Upper endpoint*		
Number	0	1	6
559	0.85	0.07	0.08
140	0.07	0.85	0.08
57	0.30	0.18	0.53

Instruments	*Upper endpoint*		
Number	0	1	6
546	0.85	0.06	0.09
116	0.06	0.73	0.21
94	0.26	0.34	0.40

Chemicals	*Upper endpoint*		
Number	0	1	6
514	0.75	0.08	0.17
128	0.06	0.53	0.41
114	0.33	0.18	0.49

Processes	*Upper endpoint*		
Number	0	1	6
521	0.76	0.05	0.19
90	0.13	0.61	0.26
145	0.23	0.29	0.48

Machinery	*Upper endpoint*		
Number	0	1	6
510	0.81	0.06	0.13
76	0.11	0.70	0.20
170	0.24	0.26	0.50

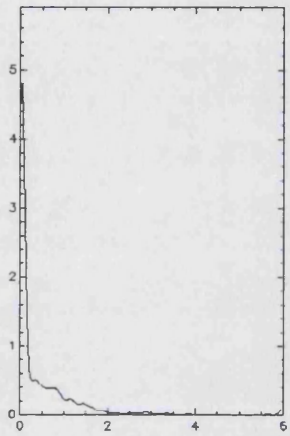
G.&C.Eng.	*Upper endpoint*		
Number	0	1	6
553	0.86	0.02	0.13
35	0.14	0.57	0.29
168	0.14	0.12	0.74

Table 2.4 Correspondence between the Fraunhofer Institute classification and the IPC.

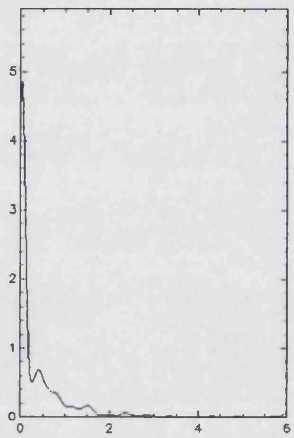
Technological Field	IPC codes
1. Electrical engineering	G05F, H01B, H01C, H01F, H01G, H01H, H01J, H01K, H01M, H01R, H01T, H05B, H05C, H05F, H05K, F21, H02
2. Audio-visual technology	G09F, G09G, G11B, H03F, H03G, H03J, H04R, H04S, H04N3, H04N5, H04N9, H04N13, H04N15, H04N17
3. Telecommunications	G08C, H01P, H01Q, H03B, H03C, H03D, H03H, H03K, H03L, H03M, H04B, H04H, H04J, H04K, H04L, H04M, H04N1, H04N7, H04N11, H04Q
4. Information technology	G06, G10L, G11C
5. Semiconductors	H01L
6. Optics	G02, G03B, G03C, G03D, G03F, G03G, G03H, H01S
7. Control and measurement technology	G01B, G01C, G01D, G01F, G01G, G01H, G01J, G01K, G01L, G01M, G01N, G01P, G01R, G01S, G01V, G01W, G04, G05B, G05D, G07, G08B, G08G, G09B, G09C, G09D, G12
8. Medical technology	A61B, A61C, A61D, A61F, A61G, A61H, A61J, A61L, A61M, A61N
9. Organic chemistry	C07C, C07D, C07F, C07H, C07J, C07K
10. Macromolecular chemistry, Polymers	C08B, C08F, C08G, C08H, C08K, C08L, C09D, C09J, C13L
11. Pharmaceuticals and Cosmetics	A61K
12. Biotechnology	C07G, C12M, C12N, C12P, C12Q, C12R, C12S
13. Materials, Metallurgy	B22, C01, C03C, C04, C21,C22
14. Food and Agriculture	A01H, A21D, A23B, A23C, A23D, A23F, A23G, A23J, A23K, A23L, C12C, C12F, C12G, C12H, C12J, C13D, C13F, C13J, C13K
15. Chemical engineering	A01N, C05, C07B, C08C, C09B, C09C, C09F, C09G, C09H, C09K, C10B, C10C, C10F, C10G, C10H, C10J, C10K, C10L, C10M, C11B-C-D
16. Surfaces	B01B, B01D, B01F, B01J, B01L, B02C, B03, B04, B05B, B06, B07, B08, F25J, F26
17. Materials processing	B05C-D, B32, C23, C25, C30
18. Thermal processes	A41H, A43D, A46D, B28, B29, B31, C03B, C08J, C14, D01, D02, D03, D04B, D04C, D04G, D04H, D05, D06B, D06C, D06G, D06H, D06J, D06L, D06M, D06P, D06Q, D21
19. Oil and Basic material chemistry	F22, F23B, F23C, F23D, F23H, F23K, F23L, F23M, F23N, F23Q, F24, F25B, F25C, F27, F28
20. Environmental technology	A62D, B01D46, B01D47, B01D49, B01D50, B01D51, B01D53, B09, C02, F01N, F23G, F23J
21. Machines, Tools	B21, B23, B24, B26D, B26F, B27, B30
22. Engines, Pumps	F01B, F01C, F01D, F01K, F01L, F01M, F01P, F02, F03, F04, F23R
23. Mechanical elements	F15, F16, F17, G05G
24. Handling	B41, B66, B67, B25J, B65B, B65C, B65D, B65F, B65G, B65H
25. Food processing	A01B, A01C, A01D, A01F, A01G, A01J, A01K, A01L, A01M, A21B, A21C, A22, A23N, A23P, B02B, C12L, C13C, C13G, C13H
26. Transport	B60, B61, B62, B63B, B63C, B63H, B63J, B64B, B64C, B64D, B64F
27. Nuclear engineering	G01T, G21, H05G, H05H
28. Space technology	B63G, B64G, C06, F41, F42
29. Consumer goods	A24, A41B, A41C, A41D, A41F, A41G, A42, A43B, A43C, A44, A45, A46B, A47, A62B, A62C, A63, B25B, B25C, B25D, B25F, B25G, B25H, B26B, B42, B43, B44, B68, D04D, D06F, D06N, D07, F25D, G10B, G10C, G10D, G10F, G10G, G10H, G10K
30. Civil engineering	E01, E02, E03, E04, E05, E06, E21

Figure 2.2 Cross-sectional distributions

Electronics

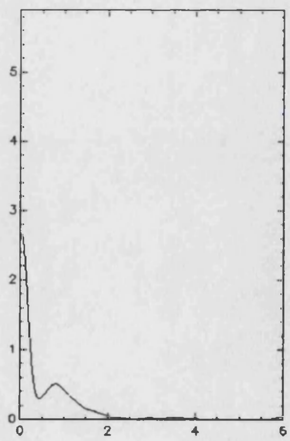


1987-1991

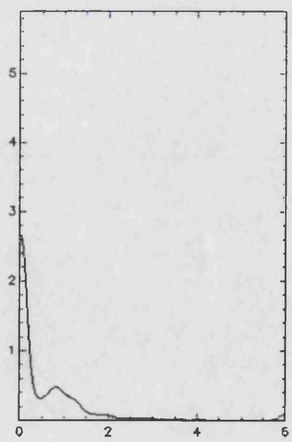


1992-1996

Instruments



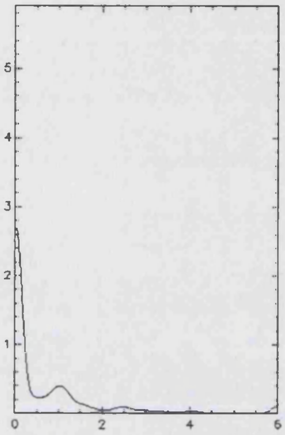
1987-1991



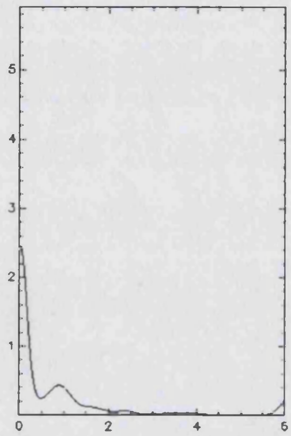
1992-1996

Figure 2.2 (cont.)

Chemicals and Pharmaceuticals

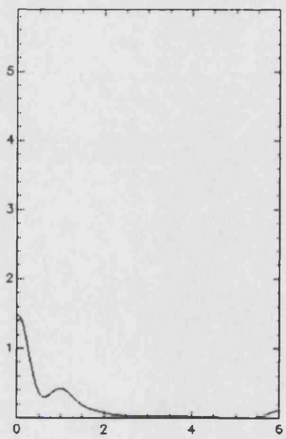


1987-1991

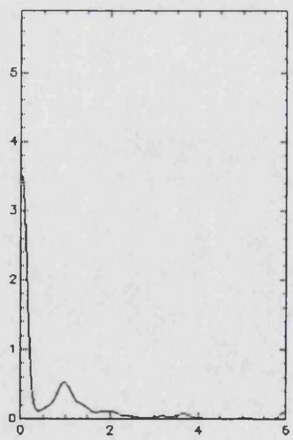


1992-1996

Processes



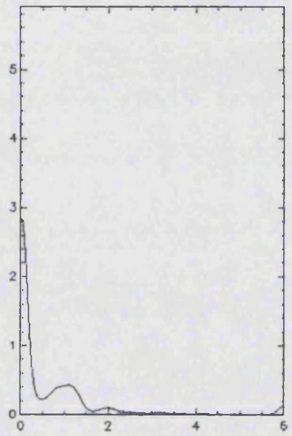
1987-1991



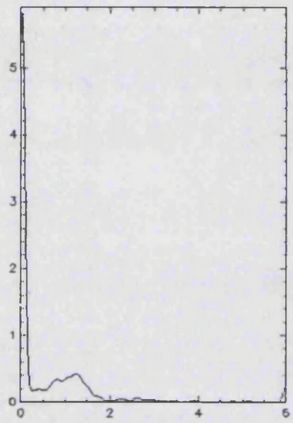
1992-1996

Figure 2.2 (cont.)

Machinery

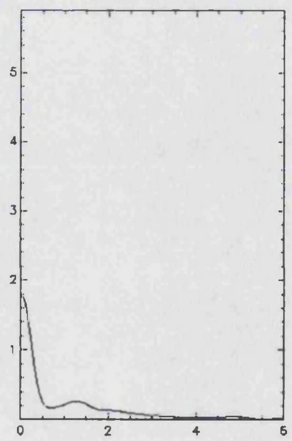


1987-1991

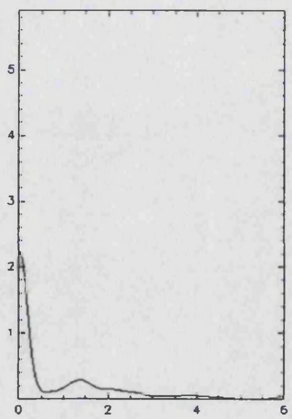


1992-1996

Consumer Goods and Civil Engineering



1987-1991



1992-1996

Figure 2.3 Markov stochastic kernels

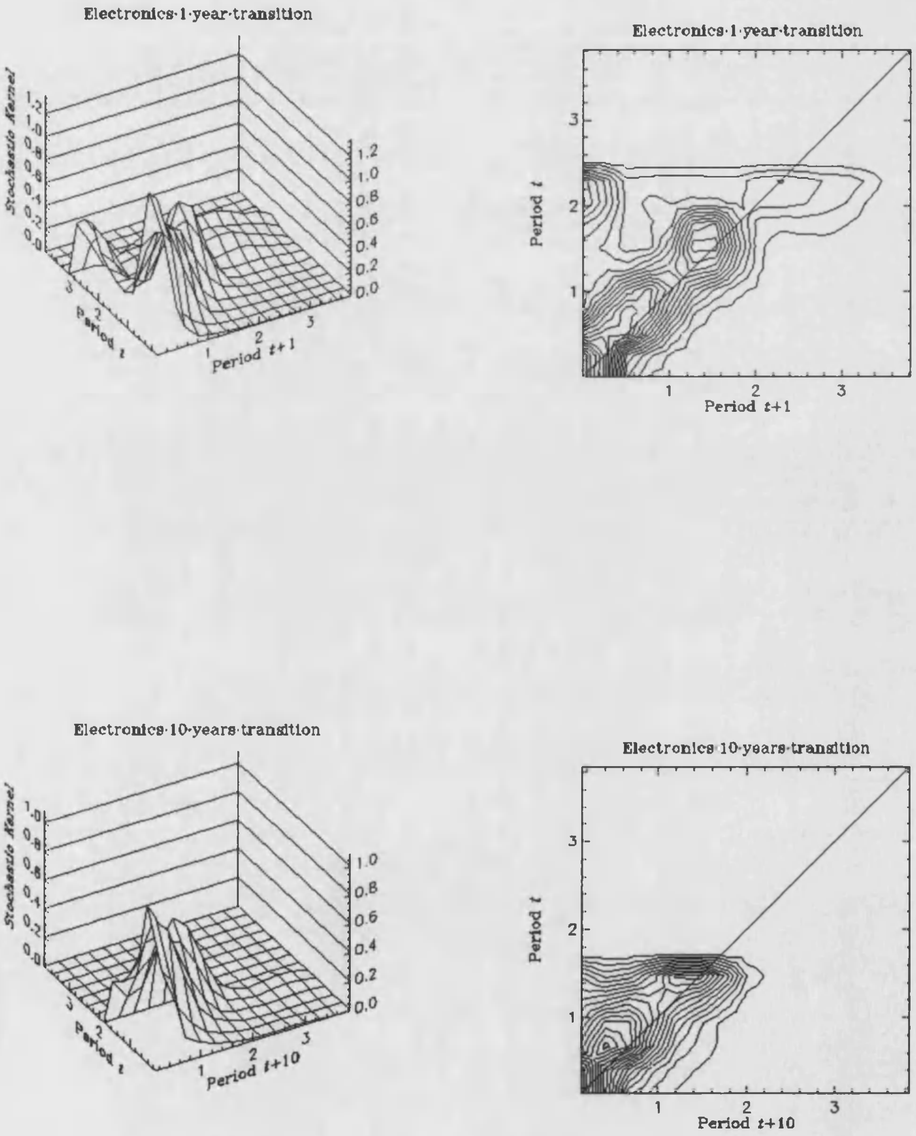


Figure 2.3 (cont.)

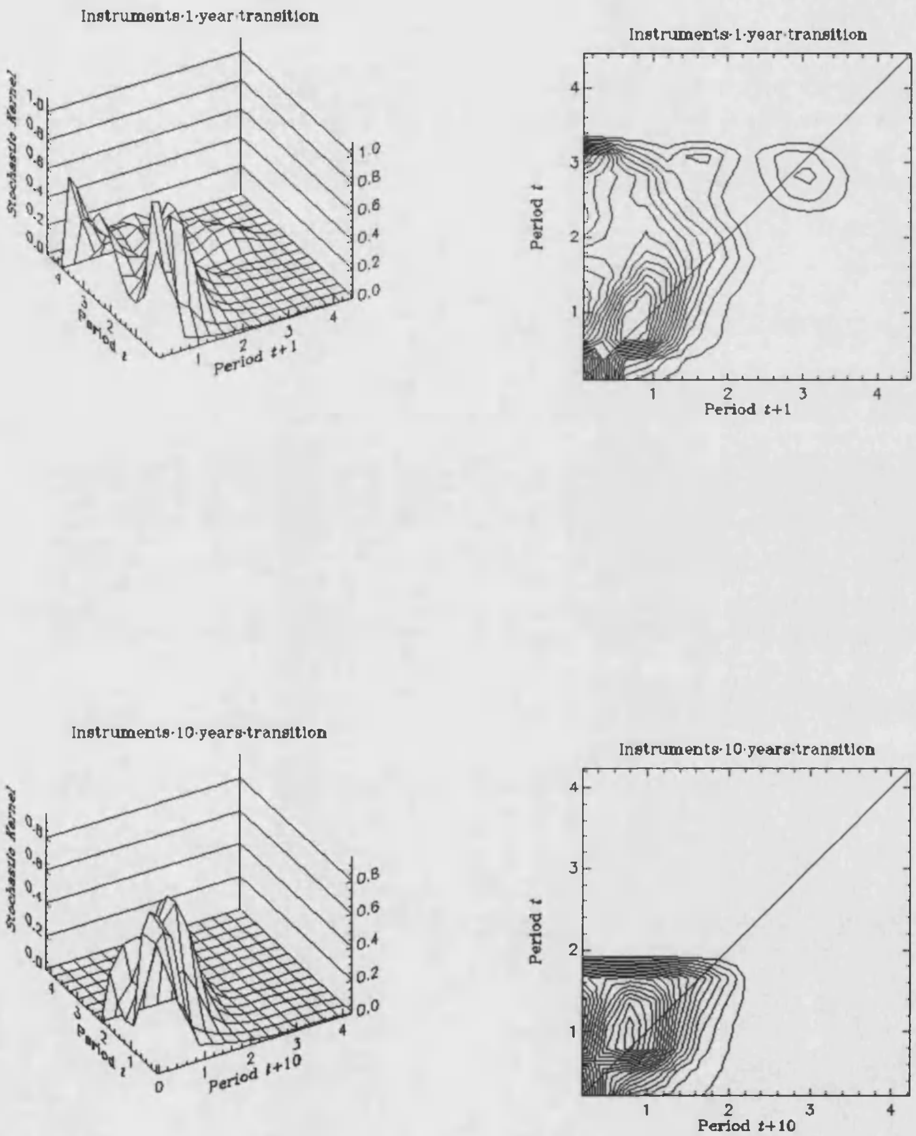


Figure 2.3 (cont.)

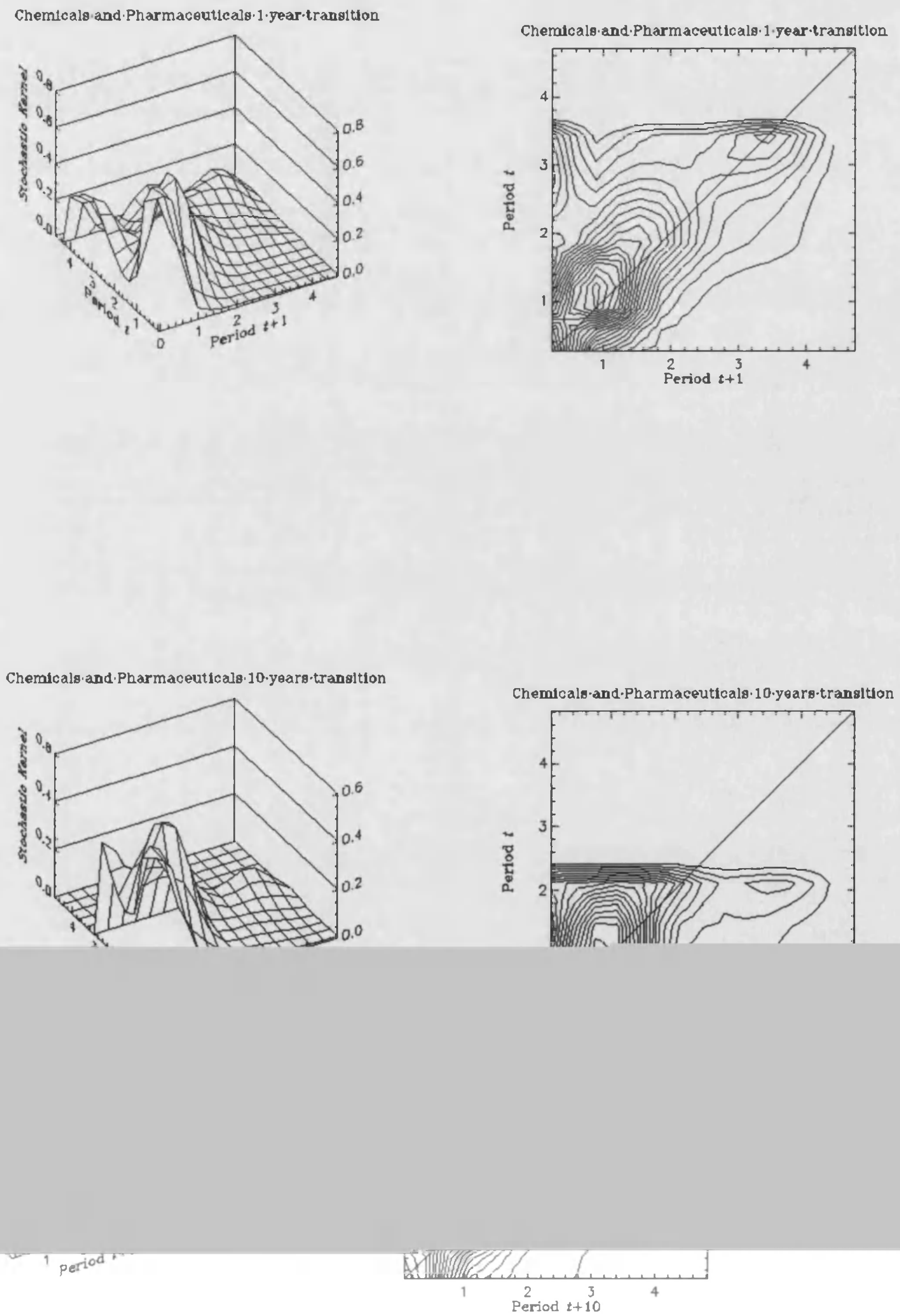


Figure 2.3 (cont.)

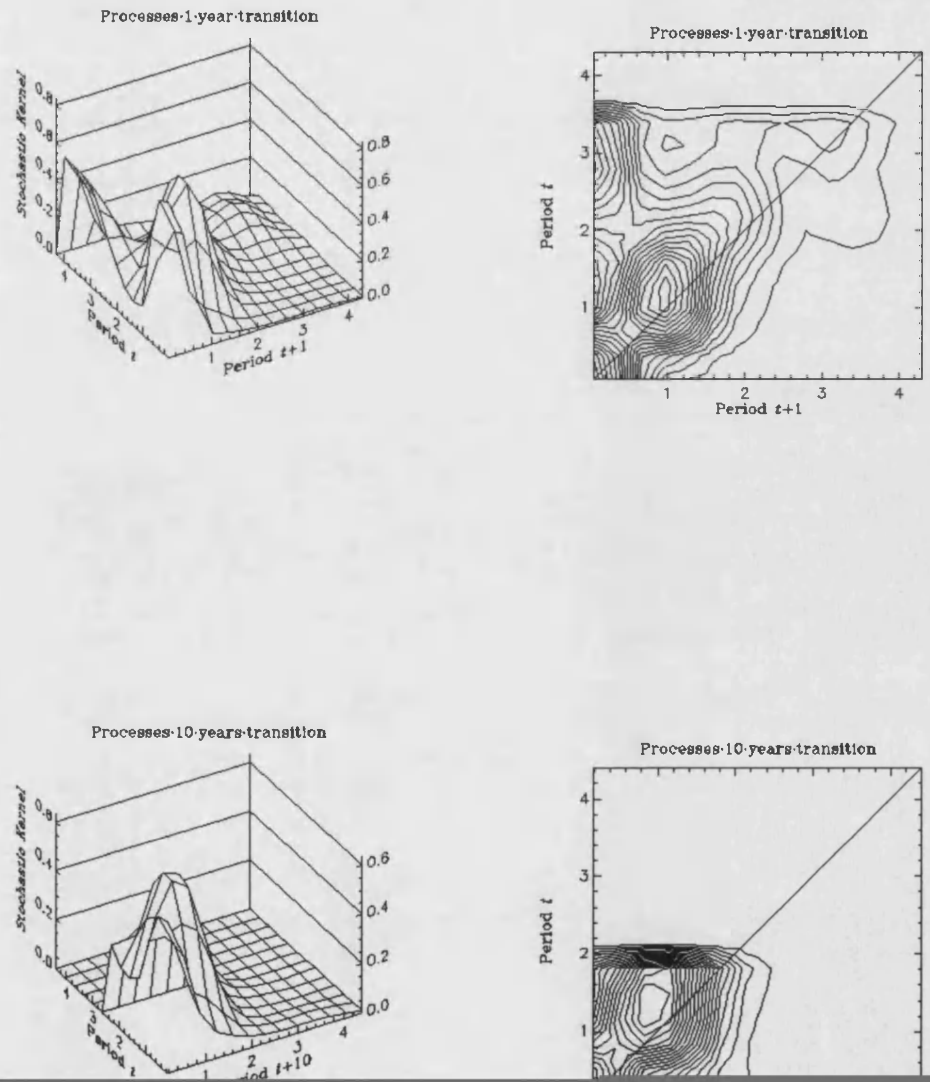


Figure 2.3 (cont.)

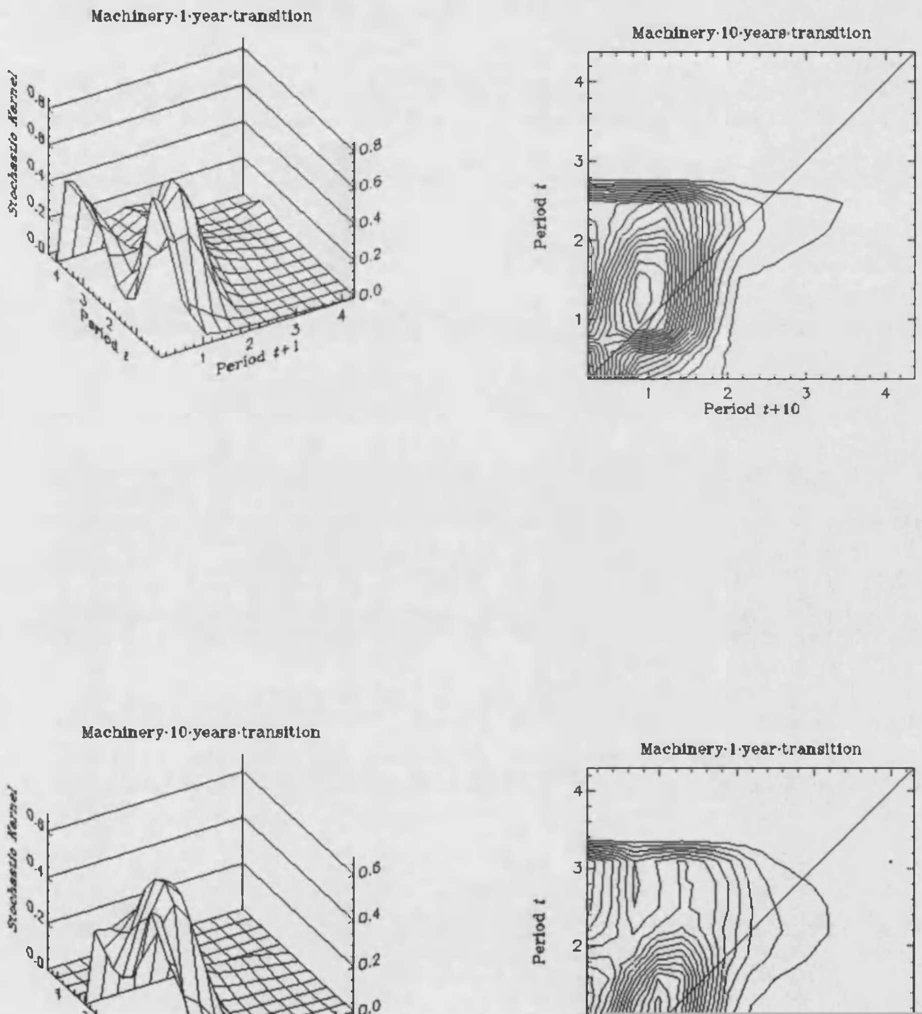
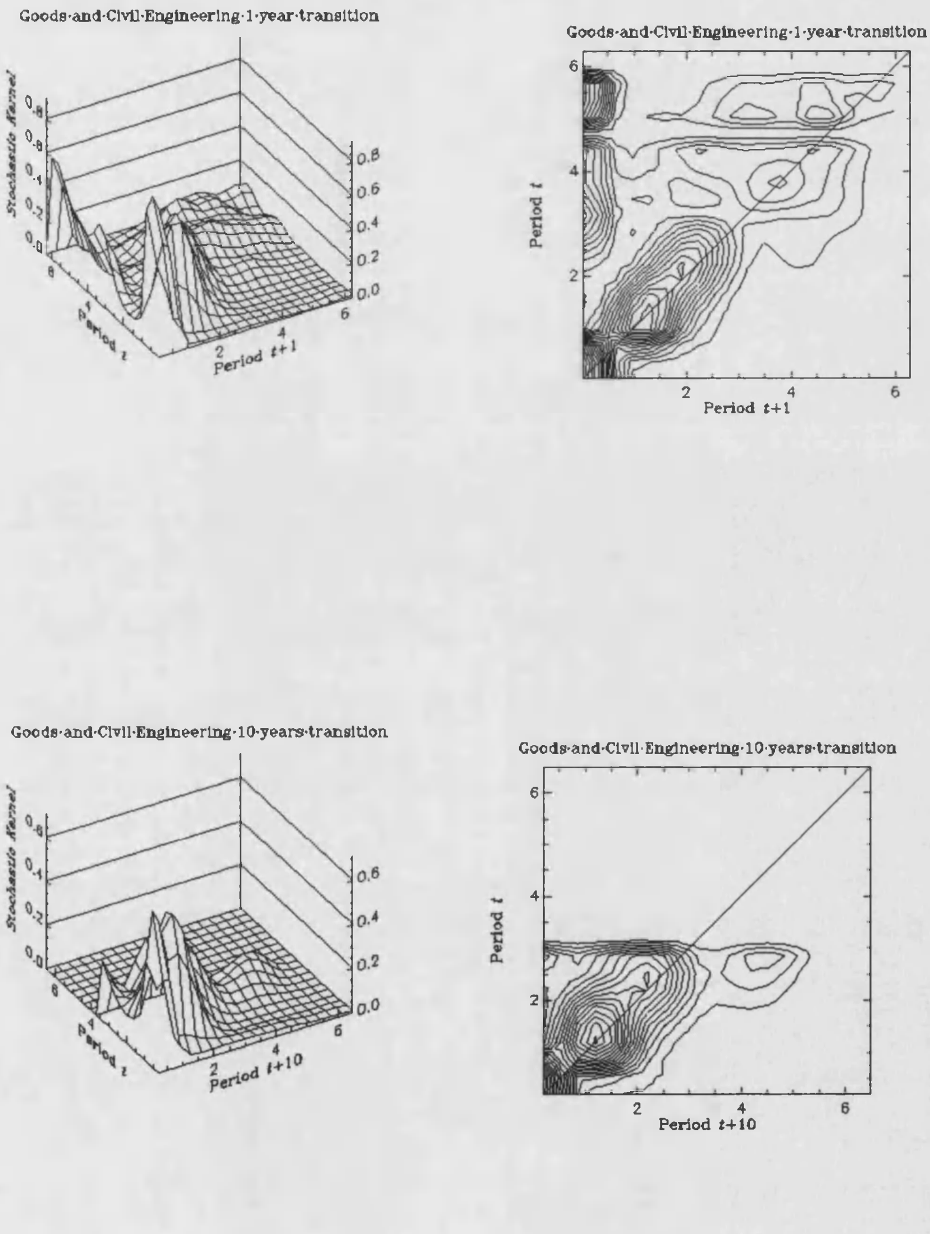


Figure 2.3 (cont.)



CHAPTER 3

TECHNOLOGY AND TRADE SPECIALIZATION DYNAMICS: A COMPARATIVE ANALYSIS

3.1 Introduction

This chapter studies the relationship between a country's pattern of specialization in trade and its pattern of specialization in technology.

The immediate motivation for this chapter lies in the contrast between the theoretical literature predicting persistence in international patterns of specialization, through either sector specific learning-by-doing (e.g. Krugman, 1987) or research and development (Grossman and Helpman, 1991, ch. 8), and the results from recent applied analyses on the tendencies towards persistence and mobility in trade specialization patterns (Stolpe, 1995; Brasili et al, 2000; Proudman and Redding, 2000). While the theoretical contributions implicitly suggest that technology and trade specialization should be closely related and evolve together, empirical analyses have been studying each of the two separately (with the only exception of Stolpe, 1995) and have undermined such theoretical contributions, finding evidence of significant mobility in trade patterns.

The origin of this mobility is still to be identified: it could be related to knowledge transfer across national borders, or to changes in countries' relative position in factor endowments. Studying the dynamics of countries' specialization in technologies, chapter 1 found no evidence of path-dependence. However, that analysis is focused on technology classes rather than industries or product groupings, as such it cannot be directly compared to the above mentioned empirical trade studies.

The first aim of this chapter is to examine directly the relationship between patterns of specialization in technology and patterns of specialization in trade. It is clear that, from a theoretical point of view, one should not expect the similarity to be too close. Patterns of trade are determined not only by differences in technology and by technological change, but also by differences in relative factor endowments, tastes, market distortions, etc. Therefore, one should expect, a priori, that the empirical similarities between the two patterns are limited. However, the question of interest is an empirical one and relates in the size of the correlation between the two specialization patterns at different aggregation levels. Empirical trade analyses are almost invariantly performed on highly aggregated industry data. If the correlation between technology and trade specialization patterns at this aggregated level is reasonably high, it will be valid to draw inferences from patterns of trade specialization in respect of mechanisms that impinge on technological specialization. If instead the correlation is low, then inferences from applied trade analyses as to underlying mechanisms such as path-dependence, which operates on technology, might be problematic.

The second aim of the chapter is to directly compare mobility in technology and trade specialization. As such, this chapter complements the analysis of chapter 1. If any mechanism of path-dependence were at work one would expect persistence in technological specialization to be at least as high as in trade specialization, if not higher.

The central theme emerging is that the correlation between technology and trade specialization patterns is extremely weak when using aggregated data, but positive and significant, although still low, when using disaggregated data. At such level of detail, technology and trade specialization levels show tendency towards convergence in the long run. Furthermore, persistence in technological specialization is significantly lower than persistence in trade specialization in the short-run, but becomes similar over a five-year horizon. The main implications are therefore that it is important to work with disaggregated data and to study the forces that lead to mobility in specialization patterns, in particular the role of international knowledge spillovers, emphasised by theory of dynamic comparative advantage and endogenous technological change.

The present analysis focuses on the five most industrialised countries. As in the previous chapters, I use patents as a measure of innovative output and map their classification (the International Patent Classification, IPC) and the commonly used Standard International Trade Classification (SITC Rev. 3) into a common classification. Technology and trade specialization patterns and dynamics can be directly compared using this correspondence. I then map the obtained classification into another commonly used one: the International Classification of Industrial Sectors (ISIC, Rev. 2) in order to asses the impact of the aggregation level on the measured relationship between technology and trade.

Following the approach adopted in the previous chapters, the analysis employs a measure derived from the well-known Balassa (1965) revealed comparative advantage index to evaluate a country's extent of trade and technology specialization in an individual sector. At any one point in time, patterns of specialization in trade and technology are each characterised by the distribution of the corresponding specialization index across sectors. The dynamics of each specialization pattern corresponds then to the evolution of the entire cross-section distribution over time.

The chapter is organised as follows. Section 2 summarises the related theoretical and empirical literature. Section 3 presents some data and measurement issues: it explains how patent and trade data have been classified according to a common classification, in order to obtain the same unit of analysis for both technology and trade specialization. Section 4 first describes the technology and trade specialization patterns of the G5 countries then looks at the correlations between technology and trade specialization profiles at different levels of aggregation. Section 4 also employs distribution dynamics techniques to estimate and compare the dynamics of technology and trade specialization patterns. Section 5 concludes.

3.2 Technology and trade

3.2.1 The treatment of technology in trade theory

Technological change was not included among the factors affecting international patterns of specialization for a long time. The classical theory of international trade focused on explanations for the pattern of trade based on Ricardo's theory of comparative advantage, which states that countries will export those goods they can produce with lowest relative costs. Ricardo saw relative labour productivity as determining differences in costs and prices and providing the basis for comparative advantage.

Neo-classical theory subsequently emphasised resource differences. The Heckscher-Ohlin-Samuelson model derives the determinants of comparative advantage in a two-good, two-factor, two-country model, predicting that a country will export the goods that use most intensively the country's more abundant factor of production. The two factors considered were capital and labour and thus the exports of a country should reflect its relative endowments of capital or labour by being relatively capital or relatively labour intensive.

Such predictions are obtained under a number of limiting assumptions: perfect competition; perfect mobility of factors within a country, but complete immobility between countries; identical demand and production function across countries. In particular, the assumption of a (static) production function common to all countries rules out technological advantage as a motivation for trade.

Many of the subsequent developments were motivated by the lack of empirical support for the factor proportions theory, starting with the well known Leontief paradox[1]. One of the roots followed by economists trying to explain away the paradox was the introduction of additional factors of production like human capital (skilled labour) or R&D expenditures. In general, results from empirical analyses pointed to the importance of both factors in explaining trade flows (Gruber et al, 1967; Keesing, 1967; Stern and Maskus, 1981). However, this approach has two fundamental weaknesses. The first relates to the endogeneity of the factors of production considered. Capital, human capital and knowledge can all be accrued over time and, as a result, cannot be considered as fixed endowments to the economy. This raises the question of the evolution of comparative advantage. The second weakness is that in the extended factor proportions approach no account is taken of the dynamic nature of technology, or the role of technology in changing the techniques available, characteristics that question the analogy between technology and labour or capital.

The role of on-going technological change (the process of innovation and diffusion) in world trade has been analysed within the technology gap theory, which originated in the informal treatment of Posner (1961) and was later formalised by Krugman (1979). In Posner's technology gap model of trade, a

[1] Leontief (1953) aimed to test the factor proportions theory for the US economy and found the famous result that the US, assumed to be the most capital-rich country in the world, was exporting more labour-intensive goods than it was importing.

country gains a temporary advantage over its trading partners through the discovery of new products and processes. For a period of time these innovations remain unique to the innovating country until they are imitated by competitors, and the innovating country loses the advantage. However, the innovating country, by having a technical superiority, can continue to innovate and maintain an advantage in a stream of new products, losing the advantage in each product, and replacing it with a new innovation. In Posner's theory, innovations are not immediately produced in countries with a cost advantage in their production, but remain in the innovating country on account of the learning period involved in the diffusion of innovation. Thus the theory predicts that countries with innovative capabilities will specialize in technology-intensive products, although, because of the changing nature of the products, the goods produced in the technology intensive country will change over time[2].

The role of imitation and diffusion is given centre place in the technology gap approach, while the enduring cumulative benefits of innovation are still ignored. A more detailed treatment of technology and its dynamic implications can be found in the neo-Schumpeterian approach (Dosi, 1988; Dosi et al. 1990), which combines the technology gap theory with a detailed view of innovation as a microeconomic process that explains how a country can maintain a cumulative advantage in the production of technology. Technology is seen as embodying specific, local, often tacit (i.e. non-codifiable), and only

[2] Vernon (1966) also provided a dynamic theory for the location of production close to the technology gap theory and generally termed the product life-cycle theory. As in the technology gap theory, Vernon postulated a country with an advantage in producing innovations. At the early stage of production of the good, production remains in the innovating country because a high level of skills is required to produce the good, its price is high and output is low. However, as the product matures and becomes standardised, price falls, production runs become longer and the production of the good moves to other countries with a cost advantage in production. The innovating country then produces another new product. The main implication of this theory is consistent with the technology gap theory: countries with high technological capabilities produce technology-intensive goods.

partly appropriable knowledge. This view is based on an evolutionary approach to technological change (Nelson and Winter, 1982) that stresses some features of the innovation process as the local nature of the search for new knowledge and techniques, and the cumulative nature of technological change. Local nature is intended in a technological sense, so that most innovations are incremental improvements on existing innovations based on past experience (Rosenberg, 1982). They are often specific to the firm in that they are based on firm-level skills and learning. At the macroeconomic level these firm-specific advantages translate into a competitive advantage for the country. Each country has a particular experience of innovation, which results from the aggregation of the innovation experiences of its firms, as well as from complementarities between different innovations, intra- and inter-industry relationships[3]. Given the existence of dynamic economies of scale in the production of knowledge, a country can build up a dynamic competitive advantage in the production of new products. This advantage can persist over time: countries can become locked-in to particular innovation and specialization patterns through their innovation history and experience (Arthur, 1989), and the nature of their institutions. This provides a microeconomic rationale for the continuation of technological differences between countries[4].

The treatment of innovation as an endogenous factor has also been pursued in the dynamic theory of comparative advantage. This strand of research (Grossman and Helpman, 1991) emphasises that technological change is an endogenous process and that comparative advantages, which determine the direction and size of international trade flows, are dynamic and evolve endogenously over time. This has brought about the issues of changes in the

[3] At the country level this pattern of innovation has been termed the national system of innovation of a country (Nelson, 1993).
[4] While this approach sees technology gaps as the most important factor motivating trade, it also stresses cost advantages in explaining the trade pattern of a country. A technology theory of trade may not be appropriate in explaining all trade flows.

countries' overall degree of specialization and the extent to which initial patterns of international specialization persist over time. The theoretical literature has reached only ambiguous conclusions concerning both these issues. If spillovers from newly generated knowledge are international in scope, then differences in technological development between countries are irrelevant for trade flows and specialization patterns. Instead, these reflect the relative factor abundance of a country and it is their accumulation which drives trade and specialization dynamics (Grossman and Helpman, 1991, ch.7). By contrast, if knowledge spillovers are geographically concentrated, and in particular if they are only effective within their country of origin, models of endogenous technological progress through sector-specific learning by doing (Krugman, 1987) or R&D (Grossman and Helpman, 1991, ch.8) predict that the evolution of a country's trade, patterns of specialization, and rate of innovation and growth depend not only on resource endowments, but also on the country's prior research experience and successes: specialization patterns will display persistence and may become locked-in over time.

The inclusion of learning, and the cumulative and endogenous nature of technology leading to endogenous comparative advantage based on differences in technology, makes dynamic models of comparative advantage close to the neo-Schumpeterian models outlined above. Technology and its accumulation are allowed to vary internationally, so that the assumption of a common production function has been dropped. As a result, differences in technology become one of the main explanations of comparative advantage.

3.2.2 Empirical analyses of trade specialization and technology

This section aims to consider some representative empirical work that has focused on the impact of innovation on trade patterns and performance. This work is highly heterogeneous in terms of the approach adopted (static vs. dynamic), the indicators used to measure technological capability and trade competitiveness, and the theoretical framework of the analysis.

Much of the empirical work is not linked to a specific theory of trade and fits into a static framework: this means taking a snapshot of the trade patterns of countries and relating them to differences in technological capabilities. Many static tests of the technology gap theory have considered correlations between a technology index and trade performance, and abstracted from other sources of trade (Soete, 1987; Amendola et al. 1998). Other tests have included additional factors as explanatory variables such as Soete (1981). He estimates a model across OECD countries for 40 industrial sectors, using an output from the innovation process (patents) in place of the more common R&D expenditures measure. He finds that technological differences between countries help explaining trade patterns for a selection of industries and that there is a great deal of heterogeneity between industries, implying that technological factors are of varying importance depending on the characteristics of the industry considered[5].

Amendola et al. (1993) test a dynamic model of the determinants of trade, estimating both the short and long-run effects of the explanatory variables, including technology, on export market shares. Using time-series data not disaggregated by sector, the authors account for the importance of past trade performance on present trade performance by including an autoregressive dependent variable in the specification. The results show the significant long-run effects of the patent and investment variables (both taken as reflecting technological capabilities), while the labour cost variable only has a short-term effect on export.

Magnier and Toujas-Bernate (1994) test a dynamic model of the impact of price and non-price factors on the export market shares of countries in

[5] The author uses a number of dependent variables, including the share of export, revealed comparative advantage (see section 3 for a definition) and the export to import ratio: the best results are obtained with the share of export as the dependent variable.

particular sectors. They include technology (relative R&D expenditure) and capacity (relative fixed investment) as indicators of non-price factors in competitiveness, along with an indication of relative prices. Using a partial adjustment specification with both country-specific and sector-specific effects, the authors find price effects to be relatively weak and non-price effects to have an important influence of trade performance in the long run.

Also Greenhalgh (1990) and Greenhalgh et al. (1994 and 1996) consider a dynamic model of price and non-price effects of trade performance over time and by sector, for a single country, the UK. The results show important long-run effects of innovation (measured as number of innovations or relative number of patents) on trade performance (measured by relative volume of export) in a significant number of sectors. However, in a number of core innovating industries trade performance does not appear to benefit from results of innovation within the sector (Greenhalgh, 1990), in high technology sectors advantages stemming from innovation are eroded in the long-run and, finally, in sectors in which world trade is dominated by multinationals innovation is not effective in supporting trade even in the short run (Greenhalgh et al 1996). Also Stolpe (1995) finds that only in a group of industries R&D activities are closely tied to production. Several of these are well-established, some even traditional industries, which rely more on gradual technological development than on revolutionary breakthroughs. By contrast, in other industries, mostly characterised by fast and radical technological change, specialization in R&D does not seem to be associated with specialization in production (Stolpe 1995, chapter D).

The paper by Harrigan (1997) specifies an empirical model of specialization consistent with the neoclassical explanation. In the model, a sector's share of GDP depends on relative factor supplies and relative technology differences (based on total factor productivity measures), and the estimated parameters have a clear connection to theoretical parameters. Using a dynamic specification, the model is estimated on panel data on manufacturing sectors in industrialised countries and relative technology levels and factor supplies are both found to be important determinants of specialization.

Another relevant work in the field is that by Gustavsson et al. (1999). The paper evaluates the impact of technology together with resource endowments, factor prices and economies of scale on international competitiveness in OECD countries. Knowledge capital stocks are obtained by cumulating R&D expenditures. Results show that competitiveness is determined not only by the R&D activity of the representative firm, but also by R&D in the domestic industry as well as economy wide stocks of knowledge, indicating the presence of local externalities. Competitiveness is also affected by factor prices and resource endowments as well as scale economies and learning by doing. Further results point to the importance of economies of scale in R&D internal to the firm. Although the authors' approach is closer to the neo-Schumpeterian literature and is static in nature, some of their results can also be related to the theory of dynamic comparative advantage.

Recall that within the theory of dynamic comparative advantage trade and specialization patterns could display hysteresis because of either sector-specific learning by doing or strong localised knowledge externalities. This has become an empirical question, on which some recent applied trade studies have been focused. Among these, Proudman and Redding (2000) study the empirical dynamics of the trade pattern of each of the G5 economies separately and claim that the degree of mobility displayed by such patterns contrasts with the results of the theoretical models of trade predicting that initial specialization patterns will become locked-in over time (e.g. Krugman, 1987; Grossman and Helpman, 1991, ch. 8). Their analysis is performed at a high level of aggregation, while Brasili et al. (2000) perform a similar exercise for both the most industrialised countries and eight fast growing Asian economies using 2-digit SITC data. They also find their evidence does not support the idea that self-enforcing mechanisms are prominent in international trade specialization.

Hence, the prediction of persistence in patterns of trade specialization is found to be at variance with the data. As the authors suggest, this finding underlines the importance of incorporating into theoretical models the economic forces which can cause changes in international specialization over time. Empirically, however, it leaves unanswered the question of which of such forces is indeed responsible of the observed mobility. International specialization patterns could change over time because technology transfer across borders is highly significant, as one strand of literature would suggest, but also because countries change their relative position in factor endowments. A recent empirical study by Redding (2002) implements a test for the factor endowment explanation and finds that mobility in trade patterns, as predicted by changes in factor endowments, is much lower than actual mobility, even after controlling for country and industry specific effects, thus suggesting a potential role for other considerations in explaining mobility, including country-specific changes in technology or relative prices.

If the hypothesis of hysteresis had any empirical content not only technological competitiveness should be positively correlated to trade competitiveness, but also persistence should be particularly pronounced in technological specialization dynamics, because that's where the factors determining hysteresis have their most direct impact. The analysis of chapter 1 found no evidence of strong reinforcing effects on the technology side, which could induce persistence in trade patterns. However, technology dynamics is there analysed with reference to technology classes, that have no straightforward correspondence to production sectors. As a consequence, it is not possible to directly compare the degree of persistence or mobility in technology and trade, as measured by the studies mentioned above.

The only study that does a direct comparison is that by Stolpe (1995). He studies the dynamics of both technology and trade specialization at a fairly aggregate level, pooling the observations of different countries together, and finds evidence of greater mobility in technology. However, the high level of aggregation does not allow him to directly compare technology and trade specialization patterns of individual countries and their evolution in time. This will be done here in the last part of section 4.

3.3 Data and measurement issues

Balassa (1965) proposed the measure of Revealed Comparative Advantage to evaluate the extent of a country's specialization in a sector. This is defined as the share of country *i* in sector *j*, relative to the country's export share in all sectors. Following Proudman and Redding (2000), the extent of an economy *i*'s trade specialization in an individual sector *j* is here characterised by the ratio of the country's share of exports in sector *j* to its average export share in all sectors:

$$RCA_{ij} = \frac{Z_{ij} / \sum_i Z_{ij}}{\frac{1}{N} \sum_j (Z_{ij} / \sum_i Z_{ij})} \tag{1}$$

where Z_{ij} denotes the value of economy *i*'s exports in sector *j* and *N* is the number of sectors.

This index normalises Balassa's measure by its cross-sectional mean at each point in time, in order to abstract from the changes in the average extent of specialization that this measure is subject to. It is analogous to the index of revealed technological advantage (*RTA*) introduced in chapter 1, which will be used here to define the extent of a country's technological specialization in an individual sector. The two indexes have the same properties, which have already been discussed in chapter 1.

At any one point in time *t*, the pattern of international specialization in trade and technology of country *i* is characterised by the distribution across sectors of *RCA* and *RTA*, respectively. The shape of each of the two density functions gives information on the overall degree of a country's specialization. If the distribution is heavily centred around unity, one can conclude that the country displays a low degree of specialization. This is the typical pattern characterising the most industrialised countries. By contrast, small and open economies are usually more specialized and internationalised as they are to some extent forced to specialize in selected niches. This translates into a distribution of the *RTA* and *RCA* indexes characterised by few fields with high values of the index and a large number of fields with very low values, that is the distribution is highly skewed with a long right tail and most of the density lying below 1.

To obtain the two specialization indexes, I use patent applications at the European Patent Office (EPO) for the period 1982-1996 and export data for the same period. Patent applications come from the EPO-CESPRI database and export data are taken from the OECD International Trade by Commodity Statistics (ITCS).

The use of patents and export data entails a classification problem that is particularly relevant for this analysis. It is primarily technical and refers to the difficulties in allocating patent data, organised by patent classes, into economically relevant industries or product groupings. In order to do a direct comparison of technology and trade specialization of a country in a sector one needs to assign both patents and export data to sectors defined by a common classification. This is not an easy task as classifications used for patents and trade data are built on different criteria, hence difficult to reconcile.

European patents are classified according to the International Patent Classification (IPC), a hierarchical classification primarily concerned with the technological characteristics of the innovation. The technical subjects refer to either the intrinsic nature of the invention, or its function or the way it is used or applied and it is difficult to allocate them into economically relevant industries or product groupings[6]. OECD-ITCS trade data are instead classified according to the Standard International Trade Classification (SITC), which is product-based and aimed at classifying all merchandise entering international trade.

There have been now different attempts to create such mappings from the technology space into the product space. The first to be developed is the Yale Technology Concordance (YTC) based on Canadian patent assignments over the period 1978-1993; it gives two-way probabilities of any IPC falling into a specific combination of industry of manufacture and sector of use, each obtained from the aggregation of SIC (Standard International Classification) codes into fifty groups[7]. A similar and very recent concordance is the OECD Technology Concordance (OTC) that maps IPC classes into ISIC (Rev. 3) sectors (Johnson, 2002). The number of sectors is larger than in the YTC, but the methodological work on the distribution of patent data by industry at the OECD is at an early stage and further improvement of the OTC is expected in the future. Finally, a concordance between 55 product fields based on the US Standard Industrial Classification and the US Patent Classification System has been developed by the Office of Technology Assessment and Forecasting at the US Department of Commerce[8].

In order to match the two classifications I use the concordance table developed by the Fraunhofer Institute for Systems and Innovation Research (FISIR), which maps the two classifications, IPC and SITC, into a third one defined by 135 highly disaggregated micro-sectors covering the chemicals,

[6] Griliches (1990) discusses this classification problem in detail.
[7] See Kortum and Putnam (1997) for a presentation and test of the YTC and visit http://www.statcan.ca/english/Subjects/Standard/sic/sic-e.htm for a full description of the SIC system used for sectoral classification in the concordance.
[8] This is the one used by Stolpe (1995).

electronics and machinery sectors[9]. I then map the highly disaggregated FISIR micro-sectors into the ISIC (Rev. 2) 22 industries[10]: this allows studying the relationship between technology and trade specialization at two different aggregation levels: one significantly higher than the other. The first is highly refined, and identifies homogeneous product groupings, thus separating high technology segments from low technology ones. The ISIC (Rev. 2) aggregation level corresponds to the level of aggregation most often adopted in trade studies and can here be used for a direct comparison with the analysis performed by Proudman and Redding (2000).

Applied trade analyses are mainly performed at a highly aggregated level, whereas industrial organisation economists argue that this might induce an aggregation problem and that 4-digit level industries would be the appropriate unit of analysis. Using the approach outlined above, it is possible to examine whether the relationship between technology and trade specialization is affected by the level of aggregation adopted. It is also possible to verify whether the results of high mobility in the empirical dynamics of trade patterns evidenced by recent studies carry through when looking at more disaggregated data.

This analysis concentrates on the five most industrialised countries (France, Germany, Japan, the United Kingdom and the United States). The data I employ cover industries belonging to the *chemical* sector (Industrial Chemicals, Pharmaceuticals and Rubber and Plastic Products), the *machinery* sector (Fabricated Metal Products, Non-electrical Machinery, Motor Vehicles, Aerospace) and the *electronics* sector (Computers and Office Machinery, Electrical Machinery, Communications Equipment and Semiconductors, Instruments). Restricting attention to the FISIR classification does not limit the

[9] Details on the mapping developed by the Fraunhofer Institute are given in the Appendix.
[10] These groupings are also used by the OECD Economic Analysis and Statistics Division for use in the Bilateral Trade Database. Further details are given in the Appendix.

scope of the analysis with reference to the technology side, but is more restrictive on the trade side. With reference to the G5 countries and the sample period, the data used in the analysis cover the 68 percent of the total number of patent applications[11], but only an average of 42 percent of the yearly export.

3.4 Trade and technology specialization in the G5 countries

In what follows I compare the trade and technology specialization patterns of each of the countries included in the sample. This will provide information on their overall degree of specialization in export and innovation activities with reference to the selected group of sectors. Since the analysis is limited to the five most industrialised countries, one should expect them to display a relatively low degree of specialization, that is the distributions of both the *RTA* and *RCA* indexes are likely to be unimodal, with most of the probability mass around the mean value of the two indexes. However, even if the two distributions appear similar, nothing can be said about the relationship between trade and technology specialization unless attention is focused from the analysis of the overall distribution to the analysis of the location of each micro-sector within such distribution. This implies looking at the value the *RTA* index has in a particular micro-sector and compare it the value the *RCA* index has in the same micro-sector. Thanks to the concordance between trade and patent classifications discussed in the previous section it is here possible to compare the two specialization indexes within the same micro-sector and to follow their evolution over time.

Before moving to the empirical analysis, it is worth reminding that the analysis faces two sample limitations. First, the *RTA* and *RCA* indexes have not

[11] There are 580621 patent applications at the EPO from the G5 countries over the period 1982-1996. Of these, 396555 belong to the chemicals, electronics and machinery sectors as defined by the FISIR classification.

been computed relative to world totals, but only relative to the totals in the five sample countries. This would not matter much if the sectoral composition of export and patenting in all countries excluded from the sample was, on average, the same as that of the countries included, but this is likely to be rather different, especially for the less developed countries. The second limitation originates from the availability of a detailed concordance between trade and patent classifications only for the chemicals, electronics and machinery sectors. As a consequence, all the other sectors are excluded from the totals used in the computation of the indexes. The computed absolute values of the two specialization indicators are therefore misleading as measures of specialization relative to the world and to the whole manufacturing sector. However, in the present context, the absolute values are of little interest compared with the relative specialization position of specific industries in the countries and its dynamics.

3.4.1 An overview of specialization patterns

As mentioned at the beginning, a country's international specialization pattern can be viewed as the distribution of its degree of specialization across micro-sectors. Table 3.1 provides some basic summary statistics describing the *RCA* and *RTA* distributions for each of the G5 countries with reference to two sub-periods at the beginning and at the end of the sample:1982-86 and 1992-96. For each country the two distributions appear rather stable over time. Furthermore, while the distributions of the two specialization indexes for the same country have a similar interquartile range, the *RTA* distribution is characterised by higher dispersion in the right tail.

The same information can be drawn from the estimated the cross-sectional distributions of the *RCA* and *RTA* indexes for each country at different time periods. Figure 3.1 reports the estimated cross-sectional distributions again with reference to the two sub-periods 1982-86 and 1992-96[12].

Table 3.1 Some basic summary statistics.

country	index	min	max	stdv	q1	q3	iqr	skew	kurt
				1982-1986					
DE	*RTA*	0.00	3.68	0.53	0.71	1.20	0.49	1.46	5.18
	RCA	0.03	2.27	0.41	0.71	1.27	0.56	0.27	0.16
FR	*RTA*	0.00	11.43	0.94	0.49	1.30	0.82	3.89	29.87
	RCA	0.07	4.15	0.62	0.59	1.26	0.67	1.91	5.61
GB	*RTA*	0.00	8.84	0.88	0.48	1.27	0.79	2.77	14.12
	RCA	0.00	4.03	0.58	0.62	1.26	0.64	1.56	4.14
JP	*RTA*	0.00	6.48	0.74	0.55	1.30	0.75	1.71	6.77
	RCA	0.02	3.30	0.63	0.53	1.34	0.81	0.80	0.66
US	*RTA*	0.00	2.64	0.41	0.76	1.24	0.48	0.11	1.29
	RCA	0.00	5.13	0.69	0.55	1.27	0.72	1.85	5.72
				1992-1996					
DE	*RTA*	0.00	4.48	0.59	0.62	1.24	0.62	1.39	4.27
	RCA	0.05	2.41	0.42	0.71	1.24	0.53	0.37	0.13
FR	*RTA*	0.00	11.55	0.94	0.46	1.30	0.84	4.62	42.53
	RCA	0.04	4.17	0.58	0.68	1.19	0.50	2.13	7.30
GB	*RTA*	0.00	9.49	0.93	0.48	1.28	0.80	3.13	16.79
	RCA	0.00	3.30	0.60	0.63	1.22	0.60	1.54	3.00
JP	*RTA*	0.00	4.14	0.60	0.62	1.35	0.74	0.74	1.75
	RCA	0.01	3.74	0.65	0.48	1.44	0.96	0.72	0.90
US	*RTA*	0.00	3.06	0.43	0.75	1.21	0.46	0.71	1.97
	RCA	0.14	3.53	0.48	0.67	1.22	0.56	1.43	3.81

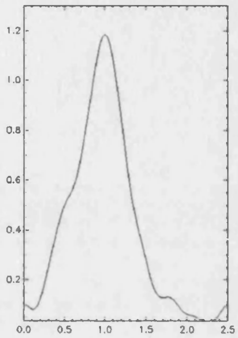
The overall degree of trade and technology specialization is never high, as the distributions are mostly centred around the mean value of the index, or slightly below it (see the UK and France, for example). In particular, the US shows a marked tendency towards decreasing trade specialization, as the peak of the distribution moves to the right, towards one, and rises in time. Note also, that the *RTA* and *RCA* distributions across sectors for the same country and in

[12] The estimation procedure is the same adopted in the previous chapters: all the densities are estimated by Gaussian kernel smoothing, taking non-negativity into account and following the procedure and automatic bandwidth choice from Silverman (1986: 2.10 and 3.4.2).

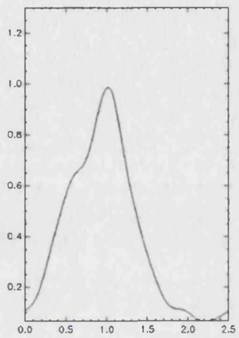
the same period can be quite different. The overall degree of specialization is higher in technology than in trade for both France and the UK (the peak of the RCA distribution is higher and dispersion is lower), while the opposite case applies in Japan. This country also shows a tendency towards increasing specialization in trade (although still centred around one, the *RCA* distribution becomes wider), while France shows the opposite tendency (the peak of the *RCA* distribution moves to the right towards one and decreases in time). Finally, the increase in spread of the *RTA* distribution for Germany might signal a tendency towards increasing overall specialization in technology[13].

Figure 3.1 Estimated cross sectional distributions

GERMANY

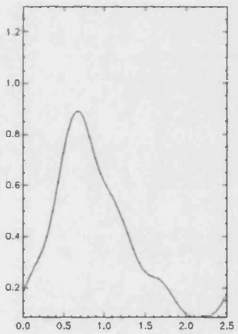


RTA: 1982-1986

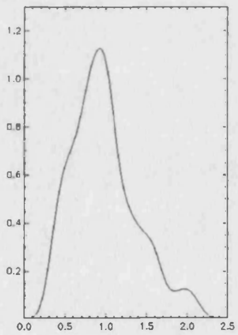


RTA: 1992-1996

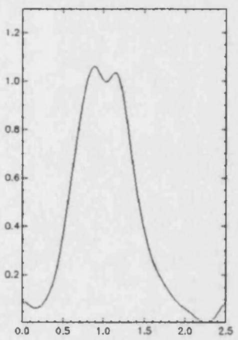
UNITED STATES



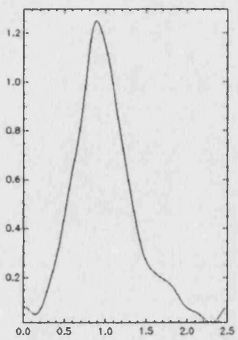
RCA: 1982-1986



RCA: 1992-1996

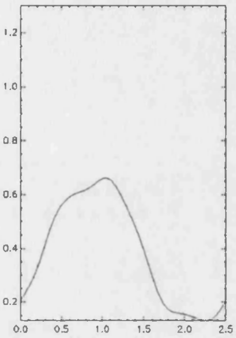


RTA: 1982-1986

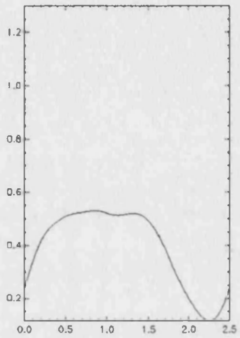


RTA: 1992-1996

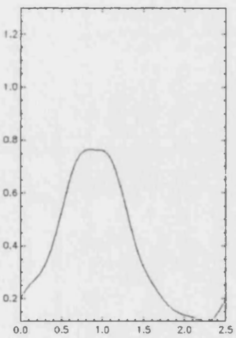
JAPAN



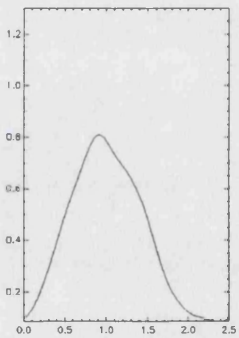
RCA: 1982-1986



RCA: 1992-1996

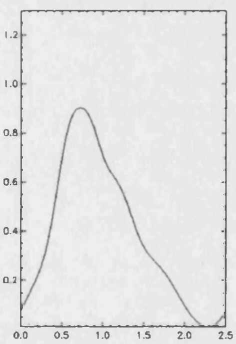


RTA: 1982-1986

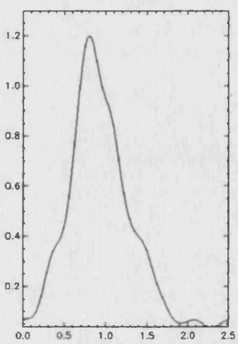


RTA: 1992-1996

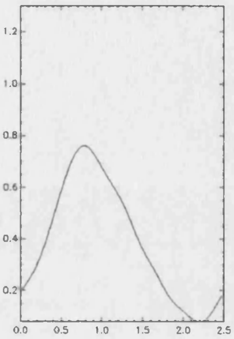
FRANCE



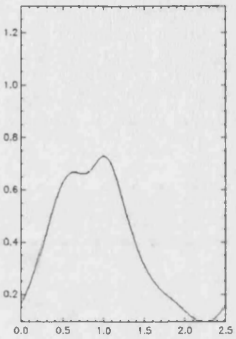
RCA: 1982-1986



RCA: 1992-1996

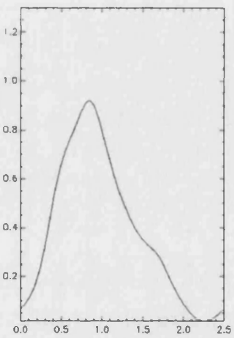


RTA: 1982-1986

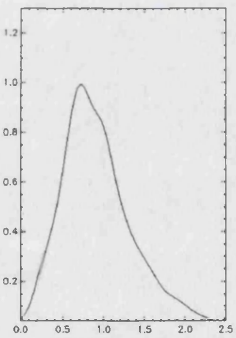


RTA: 1992-1996

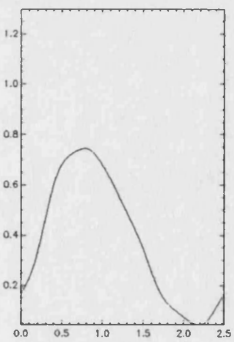
UNITED KINGDOM



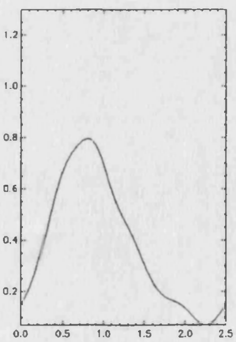
RCA: 1982-1986



RCA: 1992-1996



RTA: 1982-1986



RTA: 1992-1996

The preliminary inspection of the two specialization patterns for each country suggests that the overall specialization in trade and technology are not necessarily close, as the shape of the *RCA* and *RTA* distributions for the same country in the same sub-period can be rather different. However, the relative position of sectors within the two distributions might be similar, so the next section will evaluate the correlation between the two indexes in the same country-micro-sector pair. Furthermore, the change in the shape of the *RCA* and *RTA* distributions between the two sub-periods does not reveal substantial differences in mobility of technology and trade specialization. However, a complete evaluation of mobility requires the analysis of intra-distribution dynamics, which is presented in section 4.3.

3.4.2 The relationship between RCA and RTA at different aggregation levels

Once a mapping for export and patents into common sectors is available, simple linear correlations between trade and technology specialization can provide a very preliminary answer about the strength of their relationship. Table 3.2 reports the sample correlations between vectors of *RTA* and *RCA* at the two levels of aggregation (ISIC Rev. 2 and FISIR) in three successive sub-periods. As the table shows, correlations between technology and trade specialization at the high aggregation level are only significant in the last sub-period (1992-1996), whereas correlations at a low level of aggregation are always positive and significant, even though not very high[14].

[14] Correlations are calculated as the sample Pearson product-moment correlation between two variables X and Y:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of x and y. Probability values are obtained by treating $\sqrt{(n-2)}r / \sqrt{(1-r^2)}$ as coming from a *t* distribution with *n*-2 degrees of freedom.

The picture emerging from the correlation table is one of no or low consistency between technology and trade specialization. However, even if patents are a very informative indicator of innovative output, they may not prove particularly useful in those sectors where they do not represent an important way to protect an innovation from imitation. Correlations between technology and trade specialization may very well be not significant in those sectors where patents are not a good indicator of technological change and be, instead, significant and high in those sectors where patents are a good proxy for innovation.

Table 3.2 Simple linear correlations between *RTA* and *RCA* at two different aggregation levels over three sub-periods.

	1982-1986	1987-1991	1992-1996
ISIC2	0.07	0.22	0.41^{***}
FISIR	0.27^{***}	0.38^{***}	0.39^{***}

Note: Figures with three stars superscripts are significant at the one percent level.

Therefore, it is more appropriate to group sectors according to their patenting intensity (i.e. number of patents per dollar value of sales) and then evaluate the correlations between *RTA* and *RCA* vectors in the different groups. One should then expect correlations to become weaker when moving from high to low patent intensity sectors. Table 3.3 shows such sample correlations. Patenting intensity has been calculated with reference to the US, as the number of US patent applications in a sector divided by total sales (output) at constant prices in the same sector, averaged over the relevant sample period. Patenting intensity in a sector is then considered to be high or low if it belongs to the first or last quartile of the sample distribution, respectively[15].

[15] See the Appendix for the list of sectors classified as low, medium and high patenting

With only two exceptions, resulting correlations are still not significantly different from zero when calculated with reference to ISIC (Rev. 2) sectors. By contrast, at the low aggregation level represented by the FISIR micro-sectors correlations are not significant in the low patenting intensity sectors, while they become significantly positive in sectors where patenting intensity is medium and high . Nevertheless, such correlations still remain low.

Table 3.3 Simple linear correlations between *RTA* and *RCA* in high, medium and low patenting intensity sectors.

Patent intensity		**1982-1986**	**1987-1991**	**1992-1996**
High	**ISIC2**	-0.34	-0.27	0.06
	FISIR	0.30***	0.43***	0.37***
Medium	**ISIC2**	0.03	0.27	0.50**
	FISIR	0.26***	0.37***	0.51***
Low	**ISIC2**	0.25	0.33	0.48*
	FISIR	0.21	0.21	0.16

Note: One, two and three stars superscripts stand for significance at the ten, five and one percent level, respectively. Sectors are classified as low/high patenting intensity sectors if they fall in the bottom/top 25 percent of the distribution, respectively. Medium patenting intensity sectors are those around the mean.

Overall, the above results show that it is quite important to work at a low aggregation level when studying the relationship between technology and trade specialization. This suggests that aggregated industries used in most existing empirical work include substantial heterogeneity, which might obscure the positive correlation between trade and technological performance (see, for example, Greenhalgh et al. 1994 and 1996). Nevertheless, in line with theoretical trade analyses, the low correlations reveal that technology, however important, is only one determinant of trade specialization: its pattern and dynamics are also guided by other determinants (most importantly, factor endowments). In what follows, I shall focus attention on the relationship

intensity sectors.

between technology and trade specialization dynamics over time.

3.4.3 Comparing mobility

The empirical analysis of chapter 1 has shown that the dynamics of technological specialization is characterised by high mobility. This appeared to be much higher than mobility in observed trade specialization patterns (Proudman and Redding, 2000). However, a direct comparison could not be done because the two studies did not use the same unit of analysis. Now that it is possible to evaluate a country's technology and trade specialization in the same individual micro-sector, it is also possible to compare the degree of persistence/mobility each of the two shows. A similar analysis has been done by Stolpe (1995), although the level of aggregation adopted in his study is very high and the analysis is cross-country. By exploiting the availability of two different aggregation levels, it is here possible to see whether the relationship between technology and trade specialization dynamics is affected by the level of aggregation. Furthermore, the highly disaggregated data can be used to verify whether the result of high mobility in trade patterns carries through when specialization dynamics is analysed at a low aggregation level.

Table 3.4 presents the estimated transition probability matrices for each of the G5 economies individually and for the pooled sample[16]. Like Proudman and Redding (2000), I have estimated four-state fractile transition probability matrices (i.e. boundaries between the cells have been chosen such that class-year observations are divided roughly equally between the grid cells); this implies that each grid cell corresponds approximately to a quartile of the distributions of the *RTA* and *RCA* indexes across sectors and over time.

[16] In pooling observationas across economies, I assume that the stochastic process governing the evolution of the *RTA* and *RCA* indexes is the same in all the G5 economies.

Table 3.4 One-year transition probability matrices

GERMANY

RCA	*Upper Endpoint*			
Number	0.720	0.990	1.270	2.525
445	0.92	0.07	0.00	0.00
451	0.08	0.80	0.12	0.01
423	0.00	0.13	0.72	0.14
436	0.00	0.01	0.14	0.84
Ergodic	0.273	0.262	0.236	0.229

RTA	*Upper Endpoint*			
Number	0.655	0.960	1.210	4.666
438	0.66	0.18	0.05	0.10
448	0.19	0.50	0.22	0.09
441	0.05	0.23	0.50	0.22
428	0.09	0.09	0.25	0.57
Ergodic	0.248	0.250	0.256	0.245

JAPAN

RCA	*Upper Endpoint*			
Number	0.530	0.950	1.360	3.777
442	0.94	0.06	0.00	0.00
444	0.07	0.81	0.11	0.00
443	0.00	0.09	0.76	0.14
426	0.00	0.00	0.14	0.86
Ergodic	0.267	0.212	0.252	0.270

RTA	*Upper Endpoint*			
Number	0.540	0.900	1.300	7.575
451	0.69	0.19	0.06	0.06
441	0.16	0.55	0.24	0.05
440	0.06	0.22	0.51	0.20
423	0.07	0.05	0.22	0.66
Ergodic	0.235	0.254	0.266	0.245

FRANCE

RCA	*Upper Endpoint*			
Number	0.620	0.880	1.250	5.707
446	0.80	0.16	0.03	0.01
439	0.15	0.68	0.15	0.01
437	0.01	0.16	0.71	0.12
433	0.01	0.02	0.12	0.84
Ergodic	0.239	0.264	0.262	0.235

RTA	*Upper Endpoint*			
Number	0.455	0.860	1.280	11.67
441	0.61	0.18	0.10	0.11
443	0.16	0.48	0.25	0.11
437	0.09	0.26	0.44	0.22
434	0.12	0.09	0.22	0.57
Ergodic	0.240	0.253	0.253	0.254

UNITED STATES

RCA	*Upper Endpoint*			
Number	0.530	0.830	1.250	5.363
455	0.86	0.11	0.01	0.01
442	0.08	0.77	0.14	0.01
426	0.00	0.10	0.79	0.11
432	0.01	0.01	0.12	0.86
Ergodic	0.159	0.233	0.323	0.285

RTA	*Upper Endpoint*			
Number	0.720	0.970	1.230	2.919
451	0.60	0.25	0.08	0.08
439	0.23	0.46	0.23	0.08
425	0.06	0.23	0.47	0.24
440	0.08	0.09	0.23	0.60
Ergodic	0.238	0.257	0.256	0.249

UNITED KINGDOM

RCA	*Upper Endpoint*			
Number	0.620	0.870	1.230	4.070
440	0.86	0.12	0.01	0.01
444	0.12	0.71	0.16	0.00
433	0.01	0.16	0.72	0.10
438	0.01	0.01	0.09	0.89
Ergodic	0.255	0.249	0.244	0.252

RTA	*Upper Endpoint*			
Number	0.430	0.820	1.240	10.03
443	0.57	0.20	0.10	0.14
439	0.19	0.42	0.28	0.12
439	0.10	0.30	0.39	0.21
434	0.12	0.11	0.24	0.52
Ergodic	0.245	0.259	0.252	0.244

G5

RCA	*Upper Endpoint*			
Number	0.610	0.900	1.280	5.707
2243	0.88	0.10	0.01	0.01
2171	0.10	0.76	0.14	0.01
2186	0.01	0.12	0.74	0.13
2175	0.01	0.02	0.12	0.86
Ergodic	0.236	0.248	0.261	0.255

RTA	*Upper Endpoint*			
Number	0.560	0.910	1.240	11.67
2206	0.63	0.19	0.08	0.10
2213	0.19	0.49	0.24	0.08
2204	0.07	0.24	0.47	0.22
2152	0.10	0.09	0.23	0.59
Ergodic	0.243	0.252	0.259	0.247

The results of Table 3.4 for the one-year transition period suggest a significantly higher degree of mobility in patterns of international technological specialization compared to that in trade patterns. In all the G5 countries entries along the main diagonal in the matrix describing one-year transitions of the *RTA* index are lower than the corresponding entries in the *RCA* matrix. This means that the probability of the *RTA* moving out of one grid cell after one year is higher than that of the corresponding *RCA*: for example, in the United States, the first probability ranges from 34 to 55 percent, while the second only from 12 to 22 percent. Note how this range is very similar to that found by Proudman and Redding (2000, p. 389) for the same country: 10 to 21 percent.

Indeed, looking at the set of transition probability matrices here obtained for the *RCA* and comparing them to the corresponding ones in Proudman and Redding's study[17] there is not a clear difference in the degree of mobility, regardless of the different level of aggregation here adopted. Table 3.5 reports mobility indexes obtained from the 4-states fractile matrices. As Proudman and Redding (2000) calculate the same indexes from their estimated matrices, these have also been included in the table.

The table confirms that mobility is always higher in the *RTA* index compared to the *RCA*: for each country and each mobility index the value calculated from the *RTA* transition probability matrix is greater than the value calculated from the *RCA* transition probability matrix. As for the *RCA*, the mobility indexes here calculated are greater than the corresponding ones in Proudman and Redding (2000), but the difference in each pair is relatively small. Hence, there does not seem to be a significant difference in the degree of mobility observed in trade data at different aggregation levels. Besides, as in

[17] I am using the same set of countries, the same index and I have estimated transition probability matrices in the same way as Proudman and Redding (2000). The data here used differ from theirs in that I am using a much lower aggregation level and a subset of the ISIC Rev. 2 sectors Proudman and Redding (2000) employ.

Proudman and Redding (2000), Japan displays the least mobility, followed by Germany and the Unites States, while the United Kingdom and France exhibit the greatest mobility. The same ordering is found with respect to increasing *RTA* mobility (only the positions of France and the UK are inverted).

Table 3.5 Mobility indices.

		μ_1	μ_2	μ_4
	RTA	0.567	0.637	0.939
Germany	*RCA*	0.210	0.184	0.527
	P&R *RCA*	0.177	0.135	0.460
	RTA	0.620	0.682	0.961
France	*RCA*	0.320	0.315	0.709
	P&R *RCA*	0.253	0.196	0.607
	RTA	0.717	0.773	0.993
UK	*RCA*	0.297	0.259	0.682
	P&R *RCA*	0.243	0.187	0.590
	RTA	0.550	0.551	0.939
Japan	*RCA*	0.193	0.171	0.492
	P&R *RCA*	0.130	0.083	0.460
	RTA	0.607	0.638	0.965
US	*RCA*	0.213	0.207	0.526
	P&R *RCA*	0.207	0.161	0.518
	RTA	0.597	0.648	0.953
G5	*RCA*	0.240	0.187	0.571
	P&R *RCA*	0.163	0.121	0.426

Note:

$\mu_1 = \frac{n - tr(M)}{n-1}$; $\mu_2 = \sum_k \pi_k \sum_l p_{kl}|k-l|$; $\mu_4 = 1 - |\det(M)|$

Rows denominated "P&R *RCA*" report the mobility indices calculated by Proudman and Redding (2000, p. 390).

Transition probability matrices also allow evaluating the degree of mobility through the range of possible values of the two specialization indexes. Looking back at Table 3.4, again in accordance with the findings of Proudman and Redding (2000), mobility in the middle of the distribution appears greater than at the extremes, as the elements along the main diagonal are smallest in the lower- and upper-intermediate grid cells. However, contrary to what stated by Proudman and Redding (2000), that does not imply that mobility is indeed greater in the middle of the distribution: both here and in their paper the intermediate grid cells are defined by a more restricted range than the top and bottom grid cells, so that it is easier for an observation in the middle of the distribution to cross the border of the grid cell it belongs to between period t and period $t+1$[18].

The final row in each panel of Table 3.4 reports the ergodic or stationary *RTA* and *RCA* distributions, that is the distributions towards which the two patterns of international specialization are evolving. The ergodic distribution is always very close to a uniform distribution as should be the case, given that the estimated matrices are fractiles. As was explained in chapter 1, a simple linear regression of the interquartile range on time can be used to verify whether any of the G5 economies shows evidence of increasing technological and/or trade specialization in a subset of micro-sectors (this would translate into a polarisation of the *RTA* and/or *RCA* distribution towards extreme values and the emergence of a bimodal distribution). The only two positive and significant coefficients are obtained when regressing the interquartile range of the *RCA* index for Japan and the interquartile range of the *RTA* index for Germany. Thus, Japan shows a tendency towards increasing specialization in trade, while Germany shows a tendency towards increasing specialization in technology.

Recall, however, that these results refer to only a fraction of the manufacturing sector, hence cannot be generalised. Nevertheless, they are particularly interesting as the fraction here considered includes numerous high-tech sectors belonging to the chemicals-pharmaceuticals and the electronics industries. Indeed, the result of increasing trade specialization for Japan

[18] This is also confirmed by the shape of the estimated stochastic kernel (see the contour plot in Figure 3), which shows no strong evidence of different degree of mobility between intermediate values of *RTA* or *RCA*, on one side, and the extreme values, on the other side.

confirms the same finding by Proudman and Redding (2000) and the tendency towards increasing technological specialization in technology for Germany also emerged from the analysis in chapter 1.

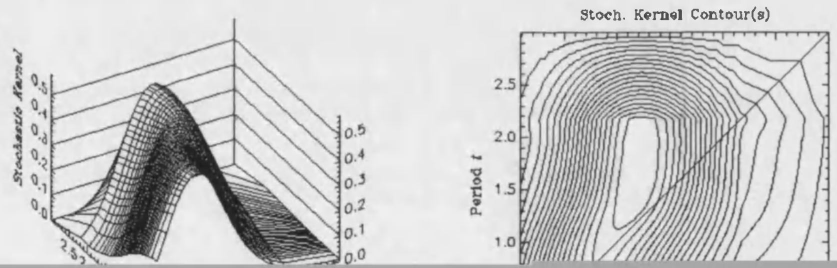
Table 3.6 Mobility indices from 5-year transition probability matrices.

		μ_1	μ_2	μ_4
Germany	*RTA*	0.613	0.688	0.964
	RCA	0.390	0.379	0.807
France	*RTA*	0.690	0.757	0.988
	RCA	0.637	0.661	0.970
UK	*RTA*	0.777	0.858	1.004
	RCA	0.523	0.520	0.920
Japan	*RTA*	0.650	0.685	0.974
	RCA	0.470	0.445	0.905
US	*RTA*	0.703	0.764	0.988
	RCA	0.567	0.568	0.949

One might now argue that mobility should be compared over a longer time horizon. This might be of particular importance as the greater mobility in *RTA* might reflect the fact that innovation occurs over long time horizons: patents may be taken out relatively infrequently, generating fluctuations in *RTA* that overstate mobility in the short run. For this reason, I have estimated five-year transition probability matrices and calculated again mobility indices. These are reported in Table 3.6. Comparing it with Table 3.5, it is clear that while mobility in technology specialization is slightly higher over five years than over just one year, mobility in trade specialization is instead much higher: *RCA* mobility indices are about the double of the corresponding ones reported in Table 3.5. Thus the difference in *RTA* and *RCA* mobility is reduced when considering longer time horizon, but this is not because the inherently random nature of the innovative process may exacerbate shifts in technology specialization, it is rather because changes in trade specialization become more frequent and wider over a longer time horizon (even if this is just a few years). As a consequence, it does not seem appropriate to think of technology as a force pushing towards persistence in trade patterns: it rather seems the opposite.

A question that arises from this result on mobility is whether it is a sign of convergence between the two specialization patterns. This can be seen by looking at mobility of trade specialization conditional on technology, rather than looking at the mobility of *RCA* and *RTA* separately.

Figure 3.2 Conditioned *RCA* dynamics for the United States.



examined. The contour plot in Figure 3.2 clearly shows that over a period of ten years *RCA* and *RTA* tend to converge one to the other. Indeed, both values of the conditioned *RCA* below one (i.e. *RCA<RTA*) and above one (i.e. *RCA>RTA*) show the tendency to move towards one: differences between *RCA* and *RTA* tend to disappear and the distribution of the conditioned *RCA* moves towards a long run ergodic distribution that gives high probability to *RCA* and *RTA* being one close to the other. The same pattern emerges from the conditioned *RCA* dynamics in the other countries. Note, however, that while this feature of specialization dynamics signals the existence of a long run relationship between technology and trade specialization, it does not say anything about the direction in which this relationship works. An empirical assessment of the relevance and structure of this relationship would require data on all the other determinants of trade specialization, which are not available at such low level of aggregation as the one here employed. This limitation is substantial as highly disaggregated units of analysis are more appropriate for studying the relationship between trade and technology specialization, as shown in the previous section.

3.5 Conclusions

This chapter has performed a first direct comparison of the degree of mobility of technology and trade specialization at the country level.

The analysis employs an index of revealed comparative advantage based on trade data to measure international trade specialization and one based on patent data to measure international technological specialization. The correlation between technology and trade specialization is found positive and significant only if a low level of aggregation is adopted. This correlation, however, remains low: the relationship between technology and trade specialization is not strong and other determinants, like factor endowments, might be of equal importance in shaping a country's trade specialization profile (see Harrigan, 1997; Gustavsson et al, 1999).

Using a low aggregation level, a distribution dynamics approach is employed to estimate and compare technology and trade specialization patterns for each of the G5 economies. A significant amount of mobility is found in both, and mobility indices reveal it to be much greater in technology compared to trade specialization over a short time horizon, while mobility is similar over a longer horizon. Therefore, the data show no evidence of reinforcing effects in the production of new knowledge that translate into persistence of technology and, consequently, trade specialization patterns. This is at odds with the strand of theory suggesting such effects (e.g. Grossman and Helpman, 1991, ch. 8) and points to the importance of studying the determinants of the observed mobility in specialization patterns. Trade theory suggests that technology might be a force towards mobility rather than persistence if there are no or few impediments to international knowledge flows (Grossman and Helpman, 1991, ch. 7): the next chapter will address this issue by studying the extent to which international knowledge spillovers affect domestic innovative performance.

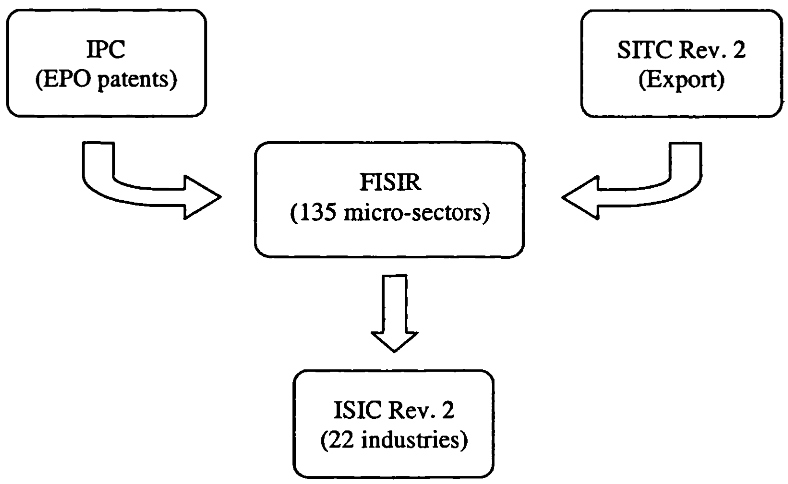
Appendix

The concordance developed by the Fraunhofer Institute for Systems and Innovation Research is between the SITC Rev. 3, a product-oriented classification, and the IPC, a technology-oriented classification. Problems arise because the latter is based on both an "application principle" and a "functional principle". A patent is assigned to an IPC class according to the following general guidelines: (i) if the object of the patent has a very specific product application, then it is classified into a technology class according to the application principle; (ii) if the innovation has a broader field of application, then the patent is assigned to multiple classes according to both the application and the functional principles; (iii) finally, if no dominant field of application exists, then the patent is assigned to a class which corresponds to the function the innovation is aimed at fulfilling, hence according to the functional principle.

While no matching problem arises in the first case, i.e. when a patented innovation is assigned to a class that corresponds to the specific product application it was developed for, it becomes more difficult to assign a patent to a sector or product group when the functional principle is involved, i.e. when the technology class the patent belongs to corresponds to a function rather than a product group. In this case, two are the roots followed by the researchers at the Fraunhofer Institute. Whenever it is possible to "split" the functional technology class into different product groupings, then the matching is done at this disaggregate level. If this is not possible, then technological and trade classes are matched according to shares. These add up to one and are chosen so to quantify the importance of the "function" for a product group, as a result of an analysis on past multiple class patents (i.e. a sort of case by case evaluation).

There is a further problem to be solved: time series for trade classified according to the SITC Rev. 3 are short as this classification was developed in 1988. OECD trade data before 1988 are only available in SITC Rev. 2, hence in order to use both patent and export data over the longest possible period I need to use the correspondence between the SITC Rev. 3 and the SITC Rev. 2[19].

Figure 3.3 The FISIR concordance between the International Patent Classification and the Standard International Trade Classification.



A further difficulty arises because in moving from revision 2 to revision 3 of the SITC some classes have been split into more product groups, with the consequence that one single SITC Rev. 2 may end up into more than one FISIR class. When this happens, I calculate the trade shares of each of the SITC Rev. 3 classes contributing to the SITC Rev. 2 one and then assign to each FISIR class involved a quota of the SITC Rev. 2 class corresponding to the share of the SITC Rev. 3 class that ends up into the FISIR class. This then allows mapping the IPC and the SITC Rev. 2 classifications into the FISIR classes,

[19] Source: Robert E. Lipsey, National Bureau of Economic Research.

thus obtaining for these a series of both patent and trade data from 1978 to 1996. Concordance between the SITC and the ISIC classifications can then be used to map the FISIR classes into ISIC Rev. 2 industries[20].

Table 3.7 ISIC Rev. 2 main sectors, as grouped by the OECD.

	ISIC Rev. 2	Description
1	31	Food, Beverages & Tobacco
2	32	Textiles, Footwear and Leather
3	33	Wood Products and Furniture
4	34	*Paper, Paper Products & Printing*
5	351+352-3522	Industrial Chemicals
6	3522	Pharmaceuticals
7	353+354	Petroleum Refineries and Products
8	355+356	Rubber and Plastic Products
9	36	Stone, Clay and Glass
10	371	Iron and Steel
11	372	Non-Ferrous Metals
12	381	Fabricated Metal Products
13	382-3825	Non-Electrical Machinery
14	3825	Office and Computing Machinery
15	3830-3832	Electric. Machinery excluding Commercial Equipment
16	3832	Radio, TV, Communication Equipment, Semiconductors
17	3841	Shipbuilding
18	3842+3844+3849	Other Transport Equipment
19	3843	Motor vehicles
20	3845	Aerospace
21	385	Instruments
22	39	Other Manufacturing

[20] For this purpose I first use the correspondence between SITC Rev. 3 and ISIC Rev. 3, and then the mapping between ISIC Rev. 3 and ISIC Rev. 2. Both can be obtained from the EU correspondence website: http://europ.eu.int/comm/eurostat/ramon.

Table 3.8 ISIC Rev. 2 sectors ordered by decreasing patenting intensity

ISIC Rev. 2	Patent Intensity
Pharmaceuticals	High
Instruments	High
Industrial Chemicals	High
Radio, TV, Communication Equipment, Semiconductors	Medium
Electric. Machinery excluding Commercial Equipment	Medium
Fabricated Metal Products	Medium
Rubber and Plastic Products	Medium
Non-Electrical Machinery	Medium
Office and Computing Machinery	Low
Motor vehicles	Low
Aerospace	Low

Table 3.9 The list of micro-sectors: Chemicals

CHEMICALS	
1) chem11	Technical polymers
2) chem12	Thermoplastics
3) chem13	Polyacetale
4) chem14	Artificial and natural caoutchouc
5) chem15	Natural polymers
6) chem16	Plastic trash
7) chem17	Plastic products
8) chem21	Inorganic chemical compounds
9) chem22	Inorganic oxygen compounds
10) chem23	Inorganic sulphide compounds
11) chem24	Other metal salts
12) chem25	Other inorganic chemical products
13) chem26	Radioactive substances
14) chem31	Synthetic textile fibres
15) chem32	Artificial textile fibres
16) chem33	Trash
17) chem41	Organic oils and fats
18) chem42	Wax
19) chem43	Artificial wax
20) chem44	Chemical products of wood or resins
21) chem51	Hydrocarbons
22) chem52	Alcohol

23) chem53	Carbon acid
24) chem54	Compounds with nitrogen function
25) chem55	Organic-inorganic compounds
26) chem56	Lactam, other heterocyclic compounds
27) chem57	Sulphamide
28) chem58	Ether, alcohol peroxide
29) chem61	Synthetic organic colours and varnishes
30) chem62	Tanning agents and paint extracts
31) chem63	Colours, varnishes, pigments
32) chem64	Glazes, sealing compounds
33) chem71	Vitamins, provitamins, antibiotics
34) chem72	Hormones and derivatives
35) chem73	Micro-organisms, vaccines
36) chem74	Reagents and diagnostics
37) chem75	Other special medicines
38) chem76	Other pharmaceutical products
39) chem77	Cosmetics (no soaps)
40) chem81	Etheric oils and perfumes
41) chem82	Soaps
42) chem83	Detergents
43) chem84	Ski-wax, furniture polishes
44) chem91	Fertilisers
45) chem92	Insecticides
46) chem101	Starch
47) chem102	Proteins
48) chem111	Explosives, gunpowder
49) chem112	Fuses, ignition chemicals
50) chem113	Pyrotechnic articles, fireworks
51) chem114	Matches
52) chem121	Additives for lubricating oil, corrosion inhibitors
53) chem122	Liquids for hydraulic brakes, anti-freezing compounds
54) chem123	Lubricants, emulsions for grease, artificial graphite emulsion
55) chem131	Gas cleansing
56) chem132	Catalysts
57) chem133	Additives for metals
58) chem134	Benzol, naphtha
59) chem135	Electronic and electro-technical chemical compounds
60) chem136	Chemical substances for constructions
61) chem137	Chemicals for fire extinguishers, liquid polychlor diphenyle

Table 3.9 (cont.). The list of micro-sectors: Electronics

ELECTRONICS:	
1) elek10	Ignition cables, electrical cars
2) elek11	Small electrical engines, electrodes
3) elek11b	Portable electrical tools
4) elek12	Motors, electrical engines and electrodes
5) elek12b	Magnetic tapes
6) elek13	Choke coils, converters, transformers
7) elek13b	Traffic lights, etc.
8) elek14	Generators and equipment
9) elek14b	Particles accelerator
10) elek15	Transformers
11) elek15b	Lasers
12) elek21	Fridges (for home and industry), air conditioning
13) elek22	Washing machines, dryers, dish washers
14) elek23	Electrical shavers, hair-cutting machines, hoovers
15) elek24	Electric heating
16) elek31	Computers and equipments
17) elek32	Computer chips and equipments
18) elek33	Photocopying machines and equipments
19) elek34	Type-writers and other office devices
20) elek41	TV, radio, TV-cameras, video-cameras, antennas, oscilloscopes
21) elek42	Microphones, loud-speakers, recorders
22) elek43	Telephones (no mobile phones)
23) elek44	Radio engineering devices
24) elek511	Circuits
25) elek512	Resistors
26) elek513	Switches, fuses
27) elek514	Control panels
28) elek521	Cables (without ignition)
29) elek522	Insulators
30) elek53	Capacitors
31) elek54	Electro-magnets
32) elek61	Electrical diagnostic devices (no X-rays)
33) elek62	X-rays
34) elek63	Instruments to show ionic beams
35) elek71	Diodes, transistors
36) elek72	Integrated circuits
37) elek8	Batteries, accumulators
38) elek9	Portable electrical lamps

Table 3.9 (cont.). The list of micro-sectors: Machinery

MACHINERY:	
1) masch10	Printing machines
2) masch11	Steam-boiler
3) masch11b	Machines for food processing
4) masch121	Steam-turbines for ships
5) masch122	Steam-turbines for steam power plants
6) masch12b	Machines to process rocks, etc.
7) masch131	Gas-turbines for aeroplanes
8) masch132	Gas-turbines for power stations
9) masch13b	Wood processing machines
10) masch14	Plastic processing
11) masch15	Cutting machine tools (saws, etc.)
12) masch16	Non cutting machine tools
13) masch17	Metal-working rolling mills
14) masch18	Soldering irons, blow lamps, welders
15) masch19	Torches, furnaces
16) masch20	Ovens, distilling apparatuses, gas distilling
17) masch21	Piston-drive engines for aeroplanes
18) masch21b	Pumps, centrifuges, filters
19) masch22	Engines for cars
20) masch22b	Conveyors
21) masch23	Engines for ships
22) masch23b	Anti-friction bearing
23) masch24	Engines for trains
24) masch24b	Valves
25) masch25	Packaging machines
26) masch26	Scales
27) masch27	Fire extinguisher, spray guns
28) masch28	Other machines
29) masch3	Water-turbines
30) masch4	Nuclear power reactors
31) masch5	Other engines
32) masch61	Agricultural machines (without tractors)
33) masch62	Tractors
34) masch7	Constructions and mining machines
35) masch8	Textile machines
36) masch9	Paper production machines

CHAPTER 4

INTERNATIONAL SPILLOVERS AND ABSORPTIVE CAPACITY: A CROSS-COUNTRY CROSS-SECTOR ANALYSIS BASED ON EUROPEAN PATENTS AND CITATIONS

4.1 Introduction

Over the last decade, the theoretical literature on growth and trade has given considerable attention to the potential role of technological externalities in generating endogenous growth and determining patterns of trade. Attention has been mainly focused on the role of international spillovers for cross-country convergence in per capita income and changes in both technological and trade specialization of countries. A growing empirical literature has addressed these issues, with contributions mainly differing along three lines, which correspond to three key questions: how do we measure knowledge spillovers? How do we assess their impact (i.e. which framework of analysis should we use)? Which level of aggregation is most appropriate for this assessment?

Knowledge external to a firm, a region or a country is obtained as a combination of R&D performed by other firms/regions/countries somehow weighted to account for the intensity of knowledge flows between the source and the destination. The measurement issue is in fact mostly related to the way such knowledge flows are inferred. Different solutions have been adopted, but since the work by Jaffe et al. (1993) patent citations have come to be considered as the most informative tool for the purpose of tracing knowledge flows.

Regardless of the way external knowledge has been measured, its impact has been assessed mainly within two different frameworks, that is by introducing the chosen measure into an aggregate production function or into a knowledge production function, which gives the relationship between newly produced knowledge (often proxied by patents) and research inputs. In the first case the aim is to assess the impact of spillovers on productivity, while in the second case their effect is measured directly on innovation. Given that one of the main difficulties in assessing the impact of knowledge spillovers lies in separating their effects from that of rent externalities (Griliches, 1979), the second approach might be preferred to the first, although this is the one that has been mostly used in the literature.

Finally, with reference to the aggregation level adopted, studies within the micro-productivity literature have mostly performed analyses at the firm level, while studies within the trade-growth literature have used a high aggregation level, with countries or regions as the unit of analysis. Therefore there is a lack of analysis performed midway between these two extremes that takes into account differences across sectors within regions or countries (thus avoiding losing relevant knowledge flows in aggregation), while still accounting for homogeneities within such sectors. This chapter takes this approach.

The impact of knowledge spillovers is here evaluated in a knowledge production function framework using data on European patents for six major industrialised countries (US, Japan, Germany, France, the UK and Italy) over the period 1981-1995. I use patent citations to trace knowledge flows within and across countries among 135 micro-sectors in the chemicals, electronics and machinery industries. Such flows are then used to estimate the effect of national versus international knowledge spillovers in the different industries and countries.

Results from different empirical studies seem to suggest that knowledge spillovers are mainly intranational rather than international in scope[1]. In one of these studies, Maurseth and Verspagen (2002) employ citations by patent applications at the European Patent Office (EPO) to trace knowledge flows across European regions: they find that patents are more likely to cite other national patents rather than foreign patents. In this chapter I show that this result arises because cross citations between European regions exclude all citations directed towards the world technological leaders (US and Japan). Once these are included in the analysis the home country effect disappears and the share of international citations is found to be particularly high in countries below the technological frontier. Consistently, international spillovers are always found to be effective in increasing innovative productivity.

The chapter then addresses a second issue, so far often neglected in the literature: the positive externality generated by international technology flows will crucially depend on the destination country's ability to understand and exploit external knowledge. Such ability is a function of the country's past experience in research, an idea analogous to the concept of *absorptive capacity* introduced by Cohen and Levinthal (1990) in the context of firms' learning and innovation.

The role of prior R&D experience in improving the ability of firms to understand and employ external knowledge has only been investigated in a few studies so far (see Griffith, Redding and Van Reenen, 2001; Griffith, Harrison and Van Reenen, 2003). The novelty here lies in the use of self-citations to measure the effect of absorptive capacity in enhancing the ability to benefit from spillovers. A self citation indicates that the firm did some research in the

[1] See, for example, Jaffe et al. (1993), Branstetter (2001), Maurseth and Verspagen (2002), Peri (2003).

past and that it has now generated a new idea building upon previous research in the same or in a related technology field. As such, self citations are a clear indication of accumulation of knowledge internal to the firm.

The empirical results show that absorptive capacity increases the elasticity of a country's innovation to both national and international spillovers. However, its effect is different depending on the position of the country with respect to the world technological frontier: the larger the gap of a country with the technological leaders, the lower is its ability to absorb and exploit external knowledge, but the larger appears its potential to increase this ability.

The chapter is organised as follows. Section 2 reviews the relevant literature on the topic. Section 3 presents the empirical model, while section 4 discusses the data used in the empirical analysis and describes some stylised facts emerging from them. Section 5 then reports the estimation results. Section 6 concludes.

4.2 Related literature

Spillovers and R&D externalities have been one of the most active areas of research in economics over the past thirty years. The reason for the still lively interest in the topic lies in their importance for growth theory and for the explanation of productivity growth. Without the social increasing returns originated by R&D externalities it is unlikely that economic growth can proceed at a constant, undiminished rate of return in the future. Moreover, the reach of spillovers has important implications for cross-country convergence in living standards. In the recent years, interest has gradually shifted to this last issue and significant research effort has been devoted in trying to assess the relevance of international spillovers and how they can be enhanced.

4.2.1 Knowledge spillovers: definition and measurement

In a pioneering paper, Griliches (1979) identifies two main sources of potential externalities generated by R&D activities: rent spillovers and pure knowledge spillovers. Rent spillovers arise when the prices of intermediate inputs purchased from other firms or countries are not fully adjusted for quality improvements resulting from R&D investment. As such, they originate from economic transactions and are the consequence of measurement "errors".

By contrast, pure knowledge spillovers arise because of the imperfect appropriability of ideas: the benefits of new knowledge accrue not only to the innovator, but "spill over" to other firms or countries, thus enriching the pool of ideas upon which subsequent innovations can be based. Hence, knowledge spillovers may occur without any economic transaction and are not the manifestation of any measurement problem.

Although the distinction between the two concepts of spillovers seems clear from the theoretical point of view, their empirical identification is far more problematic. One reason for this ambiguity is that economic transactions that originate rent spillovers may also imply some knowledge transfer[2]. Further difficulties arise because innovation by competitors may also generate strategic effects. If technological rivalry is strong and means of appropriation are effective (e.g. the scope of patent protection is wide), firms might find themselves engaged into a race for the appropriation of new profitable ideas (*patent race*). As a consequence, the positive technological externality arising from other firms' research can potentially be confounded with a negative affect due to competition[3].

Notwithstanding these difficulties, the widespread interest in the economic

[2] Together with transactions in intermediate inputs, Cincera and Van Pottelsberghe de la Potterie (2001) identify two more channels through which rent spillovers potentially operate: transactions in investment goods and the use by one firm/country of patents granted to other firms/countries. This last channel is most likely to carry knowledge spillovers as well.

[3] Jaffe (1986) and Brandstetter (2001) have found evidence of this negative effect.

implications of the existence, the magnitude and the reach of knowledge spillovers has spurred a large empirical literature. Authors have followed various approaches in the attempt to estimate the effect of spillovers. The most widely used has been to introduce a measure of the potential pool of external knowledge into a standard production function framework, either at the firm or at a more aggregate (industry, region, country) level, with the ultimate aim to assess the impact of accessible external R&D on total factor productivity (TFP). However, difficulties in measuring prices precisely and adjusting them for quality improvements make this approach not particularly suited to distinguish technological externalities from pecuniary externalities.

For this reason, some authors have adopted the *knowledge production function* (KPF) methodological framework initiated by Pakes and Griliches (1984)[4]. Within this framework research efforts and knowledge spillovers are mapped into knowledge increments, most often proxied by patents. Since the production of innovation (patents) does not require intermediate inputs and is not evaluated using prices, but simply the quantity of innovations, it minimises the role of rent externalities.

Both frameworks rely on the assumption that knowledge externalities are realised into two steps[5]. Knowledge flows represent the first step and take place whenever ideas generated by a firm/country are learned by another firm/country. Such learning creates a pool of accessible external knowledge, which then has a positive impact on productivity, however measured (this is the second step). A key issue in the empirical analyses on knowledge spillovers is then the measurement of the pool of external knowledge. This is usually built as the amount of R&D conducted elsewhere weighted by some measure of proximity in the technological or geographical space, taken to be representative

---

[4] Brandstetter (2001), Bottazzi and Peri (2003) and Peri (2003) are some of the most recent applications of this framework.
[5] Peri (2003) makes this distinction very clear.

of the intensity of knowledge flows between the source and the recipient of spillovers.

Different proximity measures have been employed in the literature. A first simple one was used by Bernstein and Nadiri (1989) who built the pool of knowledge external to a firm as the unweighted sum of the R&D spending by other firms in the same industry. This measure is fairly unsatisfactory as it assumes that a firm equally benefits from R&D of all other firms in the same industry and does not benefit at all from R&D conducted by firms in other industries. Results on spillovers based on industry measures like this might also capture spurious effects due to common industry trends and shocks.

A more sophisticated and commonly used measure of technological proximity was first introduced by Jaffe (1986). Each firm is associated to a vector describing the distribution of its patents across technology classes or its R&D spending across product fields. Such vector represents the firm's location in a multi-dimensional technology space. Proximity between two firms is then obtained as the uncentred correlation coefficient between the corresponding location vectors.

Although this measure is less likely to be contaminated by pecuniary externalities and common industry effects, evidence of its positive effect on productivity may still be unrelated to knowledge spillovers, but rather be the result of "spatially correlated technological opportunities" (Griliches, 1996)[6]. In trying to overcome these problems the most recent studies have been using the new and potentially rich source of information represented by patent citations.

Patent documents also include references to previous patents (i.e. citations) with the fundamental legal purpose to indicate which part of the knowledge

[6] Technological proximity is likely to be correlated with exogenous technological opportunity conditions. If new opportunities exogenously arise in a technological area, firms active in that area will all increase their R&D spending and improve their productivity. This would erroneously show up as a spillover effect.

described in the patent is actually claimed in the patent and which parts have been claimed by earlier patents. However, following Jaffe et al. (1993), citations can be taken as a paper trail of knowledge flows: a reference to a previous patent indicates that the knowledge of that patent was in some way useful for developing the new knowledge described in the citing patent.

For this reason, citations provide the opportunity to avoid relying on ad hoc proximity measures and look directly at the process of knowledge diffusion. Maurseth and Verspagen (2002) use citations by European patents to obtain estimates knowledge flows across European regions. Peri (2003) does a similar exercise using data on a panel of European and North American regions and then uses the obtained estimates to build a measure of accessible external R&D and assess the impact of spillovers within and across regions.

4.2.2 International knowledge spillovers

Over the last few years a great attention has been devoted to estimating the importance of international knowledge spillovers[7]. From the theoretical point of view, the interest in the reach of knowledge externalities lies in their implications for endogenous growth, trade and convergence.

If barriers to knowledge flows exist, then regions or countries' knowledge stocks may accumulate in proportion to local industrial and research activity. Increasing returns resulting from spillovers are then bounded within geographical limits and cross-country differences in levels of per capita income and in trade patterns will be persistent. By contrast, perfect technology diffusion favours the convergence of per capita output levels and leaves factor endowments as the sole determinants of trade patterns (Grossman and Helpman, 1991).

[7] A detailed survey on the topic can be found in Cincera and Van Pottelsberghe de la Potterie (2001).

The most influential contribution in the empirical literature on the topic has been the paper by Coe and Helpman (1995). They use country level data on trade shares as a proxy for the intensity of knowledge flows between countries and find that international spillovers from foreign R&D positively affect productivity growth and that this effect is larger for small countries. The previous discussion on rent spillovers should make clear why several authors have questioned Coe and Helpman's methodology to infer flows of knowledge from flows of goods. In particular, Keller (1998) provides econometric evidence that casts doubt on the effectiveness of trade as a mechanism for knowledge transfer, finding higher coefficients on foreign R&D when using random weightings instead of those used by Coe and Helpman (1995), based on trade shares.

Eaton and Kortum (1996, 1999) pursue a different line of research and derive a formal model of technology diffusion. They identify knowledge flows through cross country patenting and find that spillovers decline with geographical distance. They also show that trade is not an important channel of technological diffusion and that a country's level of education plays a significant role in the ability to absorb foreign ideas.

In a recent contribution, Bottazzi and Peri (2003) use European patent and R&D data to estimate a knowledge production function on a cross-section of European regions. They use a measure of proximity based on the geographical distance to weight R&D external to a region and find that spillovers are localised and exist only within a distance of 300 km.

Brandstetter (2001) casts doubt on the usefulness of econometric work performed at such a high level of aggregation: results obtained in such a setting are likely to reflect common demand or input price shocks or a common time trend and obscure any effect of knowledge spillovers. He argues that within countries and even within 2-digit industries there is considerable technological heterogeneity and hence performs his analysis using data on a panel of firms from US and Japan. He estimates of the impact of national and international spillovers within a knowledge production function framework, using Jaffe's uncentred correlation coefficient as a proximity measure. His results show that spillovers are more intranational than international in scope, though Japanese firms appear to benefit from the R&D of US firms to some extent.

Among the first papers to employ patent citations to study the issue of cross-border mobility of knowledge, Jaffe et al. (1993) and Jaffe and Trajtenberg (2002, chapter 7) find that a patent is typically 30 to 80 percent more likely to cite other patents whose inventors reside in the same country, than patents from other countries. This suggests that cross-border mobility of knowledge is limited and that knowledge spillovers are localised.

Maurseth and Verspagen (2002) use citations between European regions to estimate the effect of geographical distance on knowledge flows. Their results indicate that geographical distance has a negative impact on knowledge flows and that this impact is substantial. They find knowledge flows to be larger within countries than between regions located in separate countries, as well as within regions sharing the same language (but not necessarily belonging to the same country). Their results also indicate that knowledge flows are industry specific and that technological specialization of regions is an important determinant for their technological interaction as spillovers producers or receivers.

In a similar study, using the NBER patent and citations data, Peri (2003) finds that only fifteen percent of average knowledge is learned outside the region of origin and only nine percent outside the country of origin. However, his results suggest that knowledge in high technology sectors (such as computers) and knowledge generated by technological leaders (top regional innovators) flow substantially farther. Further, compared to trade flows knowledge flows reach much farther and external accessible knowledge is found to have a strong impact on innovation as measured by patent counts.

In concluding this section, I note that other authors have followed alternative approaches to the measurement of knowledge spillovers. Some works have used flows of foreign direct investment (FDI) to proxy for knowledge flows. Since FDI implies movement of capital and know-how, it has long been considered a mean of knowledge transfer and several studies find that FDI does indeed facilitate spillovers.

4.2.3 Benefiting from spillovers: the role of absorptive capacity

Recent research has started to be concerned with the ability of firms and countries to benefit from spillovers. The presumption is that firms and countries can understand external knowledge and build upon it only if they have a sufficient level of prior own knowledge and research experience.

> "*A critical component of the requisite absorptive capacity for certain types of information, such as those associated with product and process innovation, is often firm specific and cannot be bought and quickly integrated into the firm. (...) Moreover, as Nelson and Winter's (1982) analysis suggests, much of the detailed knowledge of organizational routines and objectives that permit a firm and its R&D labs to function is tacit. As a consequence, such critical complementary knowledge is acquired only through experience within the firm*" (Cohen and Levinthal, 1990, p. 135).

Along these lines, some recent papers have started to investigate the role of prior R&D experience in improving the ability of firms to understand and employ external knowledge. This issue deserves attention because if spillovers do have the potential to improve a country's growth performance, then it is important to understand the mechanisms by which they can be enhanced and made more effective.

Findings on the relevance of the absorptive capacity argument have so far been controversial. Griffith et al. (2001) use a panel of industries across twelve OECD countries to investigate whether domestic R&D, in addition to stimulating innovation, also enhances knowledge spillovers and find that domestic R&D does facilitate technology catch-up.

More recently, Griffith et al. (2003) use a sample of UK manufacturing firms to examine the role of knowledge spillovers associated with technology sourcing. They include measures of domestic and foreign external knowledge stock into the firm level production function and allow the elasticity of value added with respect to these stocks to depend on a measure of absorptive capacity and a measure of the geographical location of firms innovative activities. Although their data do not allow them to distinguish between the absorptive capacity effect and the technology sourcing effect, their results seem to suggest the latter to be more likely to affect spillovers, while the absorptive capacity effect appears quite weak.

4.3 The empirical model

I assume that in country h firms operating in micro-sector i produce new knowledge using both their own R&D and external knowledge originated either elsewhere in the same country or in another country. This idea is embodied into a production function of innovation or new knowledge:

$$Q_{hit} = f(R_{iht}, NS_{iht}, IS_{iht}, \theta, \nu_{iht}) \qquad (1)$$

where Q_{iht} is some latent measure of new technological output in micro-sector i, country h at period t, R_{iht} measures the corresponding R&D investment, NS_{hit}

is the domestic spillover pool, IS_{iht} is the foreign spillover pool and θ is the vector of unknown technology parameters.

I assume that the knowledge production function above is a Cobb-Douglas

$$Q_{iht} = R_{iht}^{\alpha} \cdot NS_{iht}^{\beta} \cdot IS_{iht}^{\gamma} \Phi_{hc} e^{\nu_{hit}} \tag{2}$$

where $\theta \equiv (\alpha, \beta, \gamma)$, ν_{hit} is an error term and Φ_{hc} captures country and industry specific effects[8] (as, for example, the set of opportunity conditions) through a set of dummy variables:

$$\Phi_{hc} = e^{\sum_{h} \delta_h D_{ih} + \sum_{c} \delta_c D_{ic}} \tag{3}$$

4.3.1 Knowledge spillovers

Estimation of equation (2) entails a series of measurement issues. The first issue relates to the measurement of the knowledge spillover variables. In the present context this involves tracing the direction and intensity of knowledge flows across micro-sectors and countries.

Knowledge flows and R&D spillovers or externalities are two distinct phases of one phenomenon, one following the other. Knowledge flows represent the first step, which takes place whenever knowledge generated by an economic agent (typically a firm) is learned by another agent elsewhere located. This diffusion process generates a stock of knowledge accessible to the recipient agent, which, through learning, then generates a positive externality on his productivity (hence the name "spillover pool"). While R&D externalities necessarily require knowledge flows to arise, knowledge flows do not automatically produce R&D externalities.

[8] Assume that micro-sector i belongs to industry c ($i \in c$).

I follow the approach initiated by Jaffe et al. (1993) and use patent citations for the purpose of tracing the direction and intensity of knowledge flows. Each patent document includes citations to previous patents that are relevant to the idea the patent is meant to protect. This establishes a close relationship between knowledge flows and patent citations: they reveal that the researchers who developed the idea knew about the ideas contained in the cited patents and that such ideas were relevant in the research process leading to the new discovery.

Unfortunately, not all the citations in a patent document are included by the inventors: some are added by the reviewers during the examination process each patent application has to go through in order to establish the novelty, originality and potential use of its content. These added citations do not necessarily reveal ideas known to the inventor. However, Jaffe et al. (1993) argue that reviewers, who are experts in a technological area, do a systematic search in that area so that this should not induce any distortion in the technological and geographical pattern of citations. Hence, I can assume that citations added by the reviewers simply add noise to the relation between knowledge flows and patent citations.

I use the information on the direction of knowledge flows implied by the pattern of citations with reference to both the technological and geographical space. For each country I consider all citations made by patents classified into each micro-sector i. I then identify the micro-sectors the cited patents belong to (i.e. their direction in the technological space) and whether they are held by other firms/institutions located in the same country (*national* citations), or by firms/institutions located in a different country (*international* citations). I also identify all citations directed to other patents held by the citing firm (*self* citations). Finally, I account for the intensity of knowledge flows using relative numbers of citations.

National spillovers are measured in the following way:

$$NS_{iht} = \prod_{j \neq i} R_{jht}^{nc_{hij}} \tag{4}$$

where nc_{hij} is the relative number of citations from patents classified into micro-sector i to patents classified into micro-sector j and held by other firms/institutions in the same country h[9]. The product is over $j \neq i$ because spillovers within the same micro-sector are already included into the own RD measure, hence their effect cannot be distinguished from that of own RD: for this reason equation (4) gives a measure of the national inter-sector pool of knowledge spillovers. Note further that this measure is obtained using only citations to *other* national firms and institutions, hence abstracting from self-citations, which cannot be regarded as a "paper trail" of knowledge flows and which account for a large proportion of overall national citations, as will be shown in the next section.

In calculating the relative number of citations I pool all citations made by patents classified in a micro-sector throughout the relevant sample period. This is equivalent to assuming constant flows for different years, an assumption which has been found to be supported by the data in a similar context (see Peri, 2003)[10].

International spillovers are measured in a similar way to national spillovers:

$$IS_{iht} = \prod_{j} FR_{jht}^{ic_{hij}} \tag{5}$$

[9] Some recent work by Peri (2003) tries to estimate the direction and intensity of knowledge flows from patterns of citations, rather than assuming that they may be represented by such patterns as I do here, along the lines of the micro-productivity literature.

[10] The advantage of this assumption is that it reduces the number of zeros in the data; the price is that of a higher serial correlation in the knowledge spillover variables, which is however a common feature in the empirical literature.

where ic_{hij} is the relative number of citations from patents applied for by firms in country h and classified into micro-sector i to patents held by firms/institutions in a different country and classified into micro-sector j[11]. FR stands for foreign R&D and is defined as:

$$FR_{jht} = \prod_{f \neq h} R_{jft}^{rc_{hf}} \tag{6}$$

where rc_{hf} is the relative number of international citations from patents held by firms in country h that are directed to patents belonging to firms or institutions resident in country f. Contrary to the national spillover measure, equation (5) includes both the international intra- and inter-sector pools of knowledge spillovers.

4.3.2 The basic specification

Substituting (4) and (5) into (2), the knowledge production function becomes:

$$Q_{iht} = R_{iht}^{\alpha} \cdot \prod_{j \neq i} R_{jht}^{\beta \cdot nc_{hij}} \prod_{j} FR_{jht}^{\gamma \cdot ic_{hij}} \Phi_{hc} e^{\nu_{iht}} \tag{7}$$

Equation (7) says that innovation in each micro-sector i in country h results from a Cobb-Douglas combination of R&D resources there used and R&D resources used in other micro-sectors and other countries. The elasticity of

[11] Note that the way I have defined national and international spillovers in (4) and (5) is less common in the microeconomic literature, where they are usually defined as a weighted average of R&D resources. The root I follow here is more common in the macroeconomic literature (see Bottazzi and Peri, 2003, for a similar application).

innovation to R&D resources other than own is then proportional to the intensity of knowledge flows between micro-sectors and countries as measured by citations.

Note that, following Branstetter (2001), I have only current own and external R&D in the knowledge production function, while one might suppose that they should enter with a long lag. With reference to own R&D, this is justified by the empirical finding that the strongest relationship between R&D and patent applications is contemporaneous (Hall, Griliches and Hausman, 1986). Furthermore, distributed lags on R&D, which is highly persistent in time, might induce a near-multicollinearity problem in the estimation[12].

Empirical research has also found evidence consistent with rapid diffusion of innovations (Caballero and Jaffe, 1993). Mansfield (1985), for example, finds that 70 percent of new product innovations "leak out" within one year and only 17 percent take more than 18 months.

There is a second measurement issue I need to deal with in order to estimate equation (7): this relates to the measurement of technological output. Since there is no direct measure of innovation I assume that some fraction of the new knowledge is patented, such that the number of new patents generated in micro-sector i, P_{iht}, is a function of its new knowledge:

$$P_{iht} = Q_{iht} e^{\sum_h \vartheta_h D_{ih} + \sum_c \vartheta_c D_{ic} + \eta_{ih}} \tag{8}$$

This is a common assumption in the knowledge production function

[12] Alternatively one could think of having a measure of R&D stock, as in Crepon and Duguet (1997). They estimate an analogous innovation function using a measure of R&D stock, built using the perpetual inventory method (see Hall and Mairesse, 1995). In this case, it can be easily shown that such measure is a linear function of current R&D. This would clearly imply a different interpretation of the coefficient on R&D, which would then be a combination of the elasticity of new knowledge to R&D, the rate of growth and the rate of depreciation of R&D.

literature[13] and in the broader innovation literature, where patents have long been considered as the best available measure of output of innovative activity. The caveats of using patent data as a measure of innovation have been widely discussed in the literature[14] and I have reported the main issues in chapter 1, where I also discuss the relevance of such critiques for this work: remarks made there are still valid in the present context.

Equation (8) controls for country specific effects and includes a set of industry dummies to account for industry-level differences in the propensity to patent, which might be related to the usefulness of patents as a tool of appropriation in industry c. Finally, equation (8) also includes individual effects, η_{ih}, to account for heterogeneity within industries and to allow for differences in the propensity to patent in each micro-sector.

Substituting (7) into (8) and taking logs I obtain my basic specification:

$$p_{iht} = \alpha \cdot r_{iht} + \beta \cdot ns_{iht} + \gamma \cdot is_{iht} + \sum_h \phi_h D_{ih} + \sum_c \phi_c D_{ic} + \eta_{ih} + \nu_{iht} \qquad (9)$$

where p_{iht} is the log of the number of patents, r_{iht} is the log of own R&D and

$$ns_{iht} = \sum_{j \neq i} nc_{hij} \ln R_{jht} \qquad (10)$$

$$is_{iht} = \sum_{i} ic_{hij} \sum_{f \neq h} rc_{hf} \ln R_{jft} \qquad (11)$$

The coefficients of the industry dummy variables in equation (9) now represent industry level differences in the propensity to patent, which are

[13] See, for example, Pakes and Griliches (1984), Branstetter (2001), Bottazzi and Peri (2003).
[14] A good reference is Griliches (1990).

functions of both the level of technological opportunity and of appropriability conditions.

I cannot directly estimate equation (9) because R&D data is not available at the same low aggregation level available for patents and citation data. R&D data is available for the 22 ISIC Rev. 2 manufacturing sectors reported in Table 4.9 in the Appendix, however given the focus of the present work on technologies in chemicals, electronics and machinery industries, only data for fifteen ISIC Rev. 2 sectors have been used as explained in the Appendix.

In order to deal with this data limitation problem, I make the following assumption:

$$R_{iht} = R^{\lambda}_{Iht}\mu_{ih} \quad \text{where } i \in I \text{ and } \mu_{ih} = e^{\xi_{ih}} \tag{12}$$

Hence, I assume that (the logarithm of) R&D expenditures within a micro-sector are a portion λ of (the logarithm of) R&D expenditures within the ISIC industry the micro-sector belongs to. This portion is assumed to be the same for all micro-sectors: differences across them are accounted for by a fixed effect component, μ_{ih}[15]. Using (12) in equation (9) the basic specification I can estimate is:

$$p_{iht} = \alpha\lambda \cdot r_{Iht} + \beta\lambda \cdot ns^{*}_{iht} + \gamma\lambda \cdot is^{*}_{iht} + \sum_{h}\phi_h D_{ih} + \sum_{c}\phi_c D_{ic} + \varepsilon_{ih} + \varepsilon_{iht} \tag{13}$$

where ns^{*}_{iht} and is^{*}_{iht} are calculated as in (10) and (11), but using the more aggregated R&D data[16].

[15] I abstract from any random time variation, given the well known relative stability of R&D expenditures over time.

[16] The individual effect in equation (13) include elements which involve summations of (weighted) individual effects components of other micro-sectors in both home and foreign

Note that the coefficients on own R&D and the spillover variables are all multiplied by λ, which is smaller than one by assumption. This should result in estimates of the elasticities that are smaller than those found in the literature[17].

It should also be mentioned that we do not observe the pure effects of knowledge spillovers on innovation by firms within a sector, which are an unambiguous positive externality. Rather, as Jaffe (1986) and, more recently, Branstetter (2001) have noticed, we observe the effects of knowledge spillovers on patents, which are not only the economic manifestation of firms' innovation, but also a tool of appropriation. If technological rivalry among the firms is intense enough and the scope of intellectual property rights is broad enough, then firms may sometimes find themselves competing for a limited number of available patents in a patent race. As a consequence, together with a positive technological externality there might be a negative effect of other firms' research due to competition. This might then result in negative estimates of the elasticities of patents to the spillover variables even though the underlying knowledge externality is positive.

In estimating equation (13) my focus will be on assessing the relevance of inter-sector and of international spillovers and on establishing differences across countries and across the three industries the data in the sample belong to: chemicals, electronics and machinery. While the idea of assessing the importance of international spillovers has received great attention in the

---

countries:

$$\varepsilon_{ih} = (\alpha\xi_{ih} + u_{ih}) + \beta\sum_{i \neq j} nc_{hij}\xi_{hj} + \gamma\sum_{j} ic_{hij}\left(\sum_{f \neq h} rc_{hf}\xi_{fj}\right)$$

Since these summations are fixed in time for each '*ih*' I can include them into an overall individual effect without loss of generality.

[17] Estimates obtained elsewhere in a similar framework (e.g. Brandstetter, 2001) are however difficult to compare to those obtained here because the micro-productivity literature has been focussing on firm-level data.

literature over recent years, studies in the field have rarely tried to evaluate the relevance of international spillovers across different sectors and the relevance of inter-sector spillovers has not been clearly assessed yet.

4.3.3 Knowledge accumulation at the firm level and absorptive capacity

The idea that knowledge generated by an economic agent flows to a different location and is learnt by some other agent crucially relies on the assumption that knowledge is, at least partially, a public good. It is however recognised that the ability to learn external knowledge often requires prior own experience. This is the well known concept of *absorptive capacity*, that is the idea that "*the more the findings in a field build upon previous findings, the more necessary is an understanding of prior research to the assimilation of subsequent findings*" (Cohen and Levinthal, 1990, p. 140).

The role of prior R&D experience in improving the ability of firms to understand and employ external knowledge has been investigated in some recent papers. While these papers examine the role of absorptive capacity in a TFP growth framework (Griffith, Redding and Van Reenen, 2001) or in a firm production function setting (Griffith, Harrison and Van Reenen, 2003) I can here directly assess its relevance on a country's innovative performance using information on self citations.

A self citation indicates that the firm did some research in the past and that it has now generated a new idea building upon previous research in the same or in a related technology field. As such, self citations are a clear indication of accumulation of knowledge internal to the firm. The higher the average number of self citations in a micro-sector the more firms operating (i.e. innovating) within such micro-sectors build upon internal knowledge in generating new ideas. If the absorptive capacity argument is correct, then such firms should also display a higher ability to understand and exploit external knowledge. A way to formalise this is to allow the elasticity of innovation (patents) to spillover pools to depend on the chosen measure of absorptive capacity. This assumption is analogous to the one made by Griffith, Harrison and Van Reenen (2003) on the elasticity of value added to the domestic and foreign external knowledge stock. In this case the aim is to assess whether the elasticity is indeed higher the more firms have been engaged into R&D activities in the same or in related technological areas.

Hence I can write the elasticity of patents to the national spillover pool (β) and their elasticity to the international spillover pool (γ) as:

$$\begin{aligned} \beta &= \beta_0 + \beta_1 \cdot self_{iht} \\ \gamma &= \gamma_0 + \gamma_1 \cdot self_{iht} \end{aligned} \tag{14}$$

where $self_{iht}$ is the number of self citations per patent in micro-sector i, in country h at time t. Differently from Griffith et al. (2003), I am not imposing the restriction that firms' absorptive capacity affects their ability to pick up domestic and foreign spillovers equally (β_1=γ_1). This is because the two spillover variables have a different "meaning": the national spillover pool here only includes inter-sector spillovers, while the international spillover variable captures the effect of both intra- and inter-sector spillovers.

Using then the expression for the elasticities to spillovers given in (14), the full specification now becomes:

$$\begin{aligned} p_{iht} = {} & \alpha\lambda \cdot r_{Iht} + \beta_0\lambda \cdot ns^*_{iht} + \gamma_0\lambda \cdot is^*_{iht} + \beta_1\lambda \cdot (ns^*_{iht} \cdot self_{iht}) \\ & + \gamma_1\lambda \cdot (is^*_{iht} \cdot self_{iht}) + \theta \cdot self_{iht} + \sum_h \phi_h D_{ih} + \sum_c \phi_c D_{ic} \\ & + \varepsilon_{ih} + \varepsilon_{iht} \end{aligned} \tag{15}$$

4.4 The data

I use patent applications[18] at the European Patent Office (EPO) and their citations, both from the EPO/CESPRI database. The analysis focuses on applications at the EPO over the period 1981-1995 by firms located in 6 countries: France, Germany, Italy, Japan, the UK and the US.

A patent document contains a detailed description of the innovation and indicates the technological class (IPC) it belongs to; it also includes the name and address of the inventor (usually one or more individuals) and of the applicant (most often a firm or an institution). Here I assign each patent to the country of residence of the applicant and consider only patent applications by firms, thus excluding individual applicants and public institutions.

I have chosen to limit the analysis to the above countries and to the 1981-1995 period because for this selected sample all firms applying for a patent at the EPO have been carefully identified and have been assigned a code. This is relevant for correctly detecting patterns of citations, as I shall later explain.

It should be noted that European patent data have been used less extensively than US patent data in the spillovers literature and that there are important differences between the two patent systems. Differently from the US Patent and Trademark Office (USPTO), the EPO acts as a single intermediary to all participating countries. Innovators may apply for a European patent up to one year after applying to their national patent office, and in most cases applications at the EPO do follow this two-stage procedure.

The national application procedure and the additional costs required to file an application at the EPO both act as a sieve that selects "good" inventions. For this reason, European patents are considered to be of higher average quality. However, the additional costs involved might induce a bias against small firms, which might then underestimate the level of localisation, if localised (national)

[18] In what follows, whenever I refer to patents, I mean patent applications.

spillovers are more important for small firms.

R&D data are taken from the OECD-ANBERD database. As already explained in the previous chapter, this entails a classification problem in that patents are classified according to the technology-based IPC classification, while R&D is classified according to the product-based ISIC classification. In order to overcome this problem, I proceed as in the previous chapter and use two different concordances: the first between the IPC and the SITC Rev. 2 (provided by Grupp-Munt, 1997), the second between the SITC Rev. 2 and the ISIC Rev.2, which I built using the OECD concordance[19].

Based on these concordances, I obtain 135 micro-sectors that represent my unit of analysis: these are the same micro-sectors employed in the analysis of chapter 3. Recall that they are analogous to product groupings and have the advantage that can be themselves grouped into three major industries: Chemicals (61 micro-sectors), Electronics (38 micro-sectors) and Machinery (36 micro-sectors). These are industries with high average R&D/sales ratio and where technological innovation is an important phenomenon, hence where it is more likely to identify the sources and effects of spillovers and of knowledge accumulation within the firm.

Table 4.1 reports the number and distribution of patents in the sample by applicant's country of residence. It shows that applications by firms in the US and Japan account for almost 60 percent of the sample. Among the European countries, Germany is the one with far the largest number of applications and a share in the sample similar to that of Japan. These shares are similar to the same countries' overall shares at the EPO[20].

Table 4.1 also shows the distribution of patents across the three main industries in the sample. Although the number of micro-sectors in the sample belonging to the chemical industry is much higher than the number of micro-

[19] See the Appendix in chaper 3 for further details.
[20] These shares for the period 1982-96 are given in chapter 1, section 5.

sectors in the electronics and in the machinery industries, its share of the total number of patents in the sample is comparable to that of the other two industries, with electronics accounting for the largest share. Indeed, the average size of a micro-sector in the electronics industry (i.e. the total number of applications over the whole sample period) is significantly larger than the average size of a micro-sector in the chemical and machinery industries.

Table 4.1 Number and distribution of patents in the sample by applicant's country of residence and by industry

Country	**Number of patents**	**% share**	**Average micro-sector size**
Germany	86228	22.6	644
France	31378	8.2	234
Italy	13411	3.5	100
Japan	87498	23.0	653
UK	26902	7.1	201
US	135587	35.6	1012
Total	381004	100	-

Industry	**Number of patents**	**% share**	**Average micro-sector size**
Chemicals	125788	33	2096
Electronics	154171	40.5	4057
Machinery	101045	26.5	2807
Total	381004	100	-

Overall, the distribution of the number of patents in each micro-sector-country pair is very skewed with a predominance of small numbers and very few large numbers, with the latter mostly belonging to the electronics industry and to either Japan or the US. Such a skewed distribution is also typical of the firm level analyses on patents.

Table 4.2 Percentage share of citations by type

Country(*)	Sector(*)	Citations		
		National(**)	International	Self
All	*All*	0.31	0.51	0.17
	Chemicals	0.29	0.50	0.21
	Electronics	0.35	0.51	0.14
	Machinery	0.28	0.56	0.16
Germany	*All*	0.25	0.56	0.19
	Chemicals	0.22	0.54	0.25
	Electronics	0.23	0.62	0.15
	Machinery	0.32	0.54	0.14
France	*All*	0.18	0.70	0.12
	Chemicals	0.18	0.68	0.14
	Electronics	0.19	0.72	0.09
	Machinery	0.18	0.70	0.12
Italy	*All*	0.13	0.74	0.13
	Chemicals	0.16	0.68	0.16
	Electronics	0.06	0.84	0.09
	Machinery	0.16	0.72	0.12
Japan	*All*	0.38	0.46	0.17
	Chemicals	0.29	0.53	0.18
	Electronics	0.44	0.41	0.15
	Machinery	0.33	0.48	0.19
UK	*All*	0.15	0.68	0.16
	Chemicals	0.18	0.63	0.20
	Electronics	0.12	0.78	0.09
	Machinery	0.14	0.71	0.15
US	*All*	0.39	0.43	0.18
	Chemicals	0.39	0.40	0.21
	Electronics	0.40	0.45	0.14
	Machinery	0.32	0.49	0.18

(*) Country and Sector refer to the citing patent.
(**) National citations are citations to national firms, universities and public research centres and exclude self citations. which are reported in the last column.

The data on citations refers to all the citations to previous European patents (i.e. patents granted by the EPO) reported in the documents of the patent applications in the sample (*backward citations*)[21]. Since each firm in the sample has been identified and has a unique code, I can separate *self* citations (i.e. citations to previous patents held by the applicant firm itself) from all *other* citations. Within these other citations I can then distinguish between citations to patents held by other national firms (*national* citations) and citations to patents held by foreign firms (*international* citations).

Table 4.2 shows the percentage distribution of national, international and self citations in different industries and countries. The table shows that the number of citations to patents held by foreign firms or public institutions is consistently higher than that of citations to national patents once one controls for self citations, the gap being particularly wide in Italy and the UK and, to a lesser extent, in France and Germany. Indeed, self citations represent an important share of overall national citations: this is equal to 35 percent in the whole sample and up to about 50 percent in Italy and in the UK.

This descriptive evidence is quite striking and does not seem to suggest the existence of significant barriers to knowledge flows across countries, rather the opposite. This is at odds with what Maurseth and Verspagen (2002) have found in a recent paper, and seems even more surprising since they also use European patent citations, although their sample only partially overlaps with mine (it includes a larger set of European countries, but excludes Japan and the US).

One reason for this disagreement could be that Maurseth and Verspagen do not have firm level data: this does not allow them to fully control for self citations, which, as shown in Table 4.1, account for a significant share of overall national citations. However, they try to mitigate the problem omitting intra-regional citations from the analysis, under the assumption that the majority of self citations should be found within the same region. Although

[21] Since I have backward citations to patents filed at the EPO and there were relatively few EPO applications in the early years there is one further reason to pool the data on citations across time when tracing knowledge flows.

some of the citations that are inter-regional may still refer to intra-firm citations, as the authors explicitly recognise, this methodology might indeed take care of a great deal of the bias self-citations generate.

There is however a second and more important reason that relates to the way the analysis by Maurseth and Verspagen (2002) is designed. They only examine flows between European regions, that is citations from one European region to another European region. In so doing they exclude citations from European regions directed towards Japan and the US, which account for the majority of patent applications at the EPO. This significantly affects the relative weight of national and international citations because a large share of the international citations of patents from European countries are directed towards Japan and the US.

With reference to my sample, this is shown in Table 4.3, which reports the directions of international citations and their relative weight. Most of the citations are to patents held by firms or institutions in the US, Japan or Germany, with the share of the first two countries ranging from 52 percent (Italy) to 69 percent (Germany). Ignoring citations directed to Japan and the US might then generate a bias in favour of national citations and induce a "border effect", as a consequence of leaving the technological leaders out of the picture[22].

Table 4.4 shows the direction of international citations for all the countries in the sample with reference to each of the three main industries. International

[22] Indeed, in the work by Peri (2003), which does not suffer from this problem, the estimate of the country border effect is significantly smaller than the one found in Maurseth and Verspagen (2002). I should however note that both the analysis by Maurseth and Verspagen (2002) and the present one sufer from the inability to control for the potentially relevant role of multinationals and their effect on international knowledge flows and on the assignment of patents to countries (I cannot control whether an innovation developed at a foreign subsidiary is patented by the home multinational). A way to partially control for this would be to assign the patent to the country of the inventor, rather than to that of the applicant firm, but in the EPO/CESPRI database information on inventors and their nationality is not yet available for all patents from the sample of countries here studied.

citations in chemicals and electronics are mostly directed towards the US. In these industries, the intensity of citations flowing towards Germany and Japan is somewhat comparable, while the UK patents appear to be cited more in chemicals than in electronics. Machinery is different in that it is German patents that receive the largest share of international citations from each of the other countries. Regardless of these differences, both Table 4.3 and Table 4.4 confirm the role of the US, Japan and Germany as technological leaders.

Table 4.3 Percentage distribution of international citations by country

		Cited country					
		DE	**FR**	**IT**	**JP**	**UK**	**US**
Citing country	**DE**	-	0.12	0.05	0.31	0.14	0.38
	FR	0.28	-	0.03	0.23	0.14	0.32
	IT	0.25	0.13	-	0.22	0.10	0.30
	JP	0.27	0.10	0.04	-	0.11	0.49
	UK	0.27	0.12	0.03	0.19	-	0.39
	US	0.28	0.12	0.04	0.39	0.17	-

Note: the percentages in the table refer to the share of citations from the citing country directed towards the cited countries (i.e. row sums are equal to 1).

Having information on the technological class of both the citing and the cited patent, I can also trace patterns of citations across micro-sectors. Although these might be thought as being narrowly defined, still about sixty percent of the citations are found to be directed to other patents classified into the same micro-sector, the percentage being slightly higher in electronics (64 percent) than in chemicals and machinery (56 percent for both)[23].

I should mention that it has elsewhere been noticed that there might exist a

[23] This pattern is consistent across countries, as can be seen in Table 4.10 in the Appendix. Note that the percentage might be higher in electronics because of the larger average micro-sector size within this industry compared to chemicals and machinery industries.

potential problem with the informative content of European patent citations. This is related to the number of citations included into the patent document by the examiners, rather than by the innovator: these citations represent knowledge not necessarily known to the innovator, hence not necessarily used in the process leading to the innovation.

This criticism is often raised in the literature and is relevant for both the European and the US patent systems, since in both cases it is patent examiners who finally determine what citations to include into a document. However, while the US system requires applicants to provide a complete description about the state of the art, the European system does not, which implies that the share of citations added by the examiners is likely to be larger in patents filed at the EPO compared to patents filed at the USPTO (Maurseth and Verspagen, 2002, p. 534). While this might increase the noise in the relation between knowledge flows and patent citations in the case of European data, it is not clear that it should lead to any specific bias.

Despite this criticism, there is little existing evidence on the validity of using patent citations as a measure of knowledge flows. A recent paper by Duguet and MacGarvie (2002) assesses the legitimacy of using European patent citations as a measure of knowledge spillovers. They use information from the CIS1 survey collected by the French Service des Statistiques Industrielles, which contains firms' responses to questions about their acquisition and dissemination of new technologies across countries. By matching firms' responses to citation counts the authors find that patent citations are indeed related to firms' statements about their acquisition of new technology. The results obtained by Duguet and MacGarvie (2002) and the analogous findings of Jaffe, Trajtenberg and Fogarty (2000) on US citations data, strengthen the case for the use of patent citations as they appear to be sufficiently correlated with knowledge flows to allow statistical analysis based on them to be informative about the underlying phenomenon of interest.

Table 4.4 Percentage distribution of international citations by country within each industry

CHEMICALS

Cited country

Citing country		DE	FR	IT	JP	UK	US
	DE	-	0.07	0.03	0.29	0.17	0.45
	FR	0.18	-	0.04	0.21	0.17	0.40
	IT	0.19	0.09	-	0.20	0.13	0.39
	JP	0.27	0.06	0.03	-	0.14	0.49
	UK	0.21	0.10	0.03	0.16	-	0.51
	US	0.28	0.10	0.03	0.33	0.25	-

ELECTRONICS

Cited country

Citing country		DE	FR	IT	JP	UK	US
	DE	-	0.12	0.03	0.35	0.09	0.40
	FR	0.22	-	0.03	0.28	0.10	0.37
	IT	0.19	0.14	-	0.28	0.08	0.32
	JP	0.19	0.10	0.02	-	0.08	0.60
	UK	0.19	0.13	0.02	0.27	-	0.40
	US	0.22	0.12	0.03	0.52	0.11	-

MACHINERY

Cited country

Citing country		DE	FR	IT	JP	UK	US
	DE	-	0.16	0.07	0.30	0.14	0.33
	FR	0.37	-	0.03	0.21	0.15	0.25
	IT	0.32	0.15	-	0.19	0.10	0.25
	JP	0.40	0.12	0.07	-	0.11	0.30
	UK	0.37	0.15	0.03	0.17	-	0.28
	US	0.37	0.16	0.05	0.27	0.15	-

Hence, although the data on R&D have some important imperfections, the data on patents and citations are very detailed and have the advantage of including the whole set of EPO patent applications and relative citations for the selected countries and industries. This allows an accurate identification of knowledge flows through citations, on which spillovers measures are based.

In the estimation, for all the countries I could not use one of the 135 micro-sectors because no clear correspondence with the R&D classification could be identified. I also dropped from the sample all the micro-sector/country pairs with zero patent counts in each year and further restricted the sample to micro-sectors/country pairs with at least fifteen patents during the sample period in order to avoid jumps due to sporadic observations.

Table 4.5 Summary statistics for the complete sample

Variable	Mean	Std. Dev.	Min	Max
Patents	35.63	67.46	0	1166
RD(*)	2626.97	2842.55	18.90	27113.57
NS(*)	42.68	357.96	0	9106.05
IS(*)	2609.35	1139.75	307.82	7494.91
self	.13	.15	0	1

(*) Units are millions of 1990 US dollars

The restrictions to the sample mainly affect the chemical industry, to which most of the micro-sectors with few patent applications belong. Hence, the final sample I use in the estimations includes 712 cross-sectional units, evenly distributed across industries (286 micro-sectors from the chemicals industry, 218 from the electronics industry and 208 from machinery industry). Table 4.5 reports the summary statistics for the selected sample.

4.5 Estimation

This section presents empirical methods and results from the estimation of equations (13) and (15).

In the estimation of both specifications the dependent variable is equal to the log of patents for micro-sector *i* in country *h* at time *t*. Since in the sample there are cross-sectional units for which the number of patents is equal to zero in some years and the logarithm of zero is undefined, I add one to all observations of the number of patents and then take the log to obtain the dependent variable used in the log-linear regressions reported below.

Although the above transformation represents the traditional and widely used procedure for dealing with this problem in the literature, there are concerns that it might bias the results. Indeed, as noted in the previous section, the distribution of patents in the sample is highly skewed, with a preponderance of small numbers and a significant percentage of zeros (this is equal to 12 percent in the complete sample). Furthermore, patents are count data and occur in integers. These characteristics are known to generate bias in the estimates of the log-linear model (see Winkelman, pp. 67-8) and motivate the estimation of alternative non-linear models.

Regardless of the model chosen (linear vs. non-linear), a concern in the estimation of both equations (13) and (15) resides in the complex structure of the individual effect, which is characterised by correlation across panels (here: country/micro-sector pairs), hence by a residual variance-covariance matrix that is no longer block-diagonal[24]. If such correlation is ignored, inferences based on OLS or random effects estimation might then be misleading since estimated standard errors are biased downward. By contrast, fixed effects estimates are conditional on the individual effects, which leaves the standard errors unaffected[25]. Furthermore, fixed effects methods ensure consistency in

[24] This is generated by the data availability problem for R&D through the presence of the spillovers variables, which are built upon it (see footnote 16).
[25] It should be noted that correlation across panels also occurs when an aggregated variable is included among the regressors (Moulton, 1986). This is the case in both specification (13) and (15), where in each time period there are repeated observations on R&D because the data

the presence of correlation between the explanatory variables and the individual effects. For the above reason, fixed effects methods, although inefficient, are to be preferred.

Before moving to the estimation models and results a final remark should be added with reference to the dependent variable. One might argue that a more appropriate measure of innovation in a field would be the count of patents weighted by the number of citations received (*forward citations*) in order to account for the quality of patents as proxy for new ideas (see Jaffe and Trajtenberg, chapter 2). This would require excluding observations belonging to the last years in the sample, effectively reducing the available period to the 1980's[26]. The benefits of this choice are however uncertain. Using US patent data and citations, Peri (2003) finds no significant difference in the estimates of the effects of R&D spillovers on innovation using weighted and unweighted patent counts. Further, as previously explained, the average quality of the EPO patents in the sample is relatively high, thus adjustment for quality through citations is unlikely to be found more significant in this setting.

The following section briefly describes the non-linear methods employed in the econometric analysis. Subsequent sections comment the empirical results presented in Table 4.6 through to Table 4.8.

4.5.1 Fixed effects non-linear regression models for count data

The basic model found in the literature to handle count data is the Poisson model, which has been extensively used to model patents as a function of R&D (see Hausman, Hall and Griliches, 1984). This model estimates the relationship

availability for such variable is limited to a higher level of aggregation than the one used for the dependent variable. The induced correlation problem is here ruled out by assumption (12), which effectively says that having aggregated R&D on the right hand side affects the size of the estimated coefficients, but not the standard errors.

[26] In the NBER data on US patents, Jaffe and colleagues found that the lag distribution of forward citations is skewed to the left, with a mode at about 3.5 years. Most of the citations are received within ten years from granting, but there can be long lags (up to thirty years).

between the arrival rate of patents and the independent variables. The dependent variable, y_{it}, is assumed to have a Poisson distribution with parameter μ_{it} which, in turn, depends on a set of exogenous variables $\boldsymbol{x}_{it}$ according to the log-linear function:

$$\ln \mu_{it} = \delta_i + \beta \mathbf{x}_{it} \tag{16}$$

where δ_i is the fixed-effect.

One way to estimate this model is to do conventional Poisson regression by maximum likelihood, including dummy variables for all individuals (less one) to directly estimate the fixed effects. If there is no specific interest in the fixed effects or if, as in this case, their number is large conditional maximum likelihood represents an alternative method[27]. Conditioning on the count total for each individual, $\sum_t y_{it}$, it yields a conditional likelihood proportional to

$$\prod_i \prod_t \left(\frac{\exp(\beta \mathbf{x}_{it})}{\sum_s \exp(\beta \mathbf{x}_{is})} \right)^{y_{it}} \tag{17}$$

which no longer includes the δ_i parameters.

The fixed effects Poisson regression model allows for unrestricted heterogeneity across individuals, but requires the mean of counts for each individual to be equal to its variance ($E(y_{it}) = V(y_{it}) = \mu_{it}$). This is an undesired feature whenever there is additional heterogeneity not accounted for by the model, i.e. when the data show evidence of overdispersion. Such problem can be dealt with by assuming that y_{it} has a negative binomial

[27] For the Poisson regression the two methods always yield identical estimates for β and the associated covariance matrix (Cameron and Trivedi, 1998), hence the choice of method is entirely dictated by computational convenience.

distribution (see Hausman, Hall and Griliches, 1984), which can be regarded as a generalisation of the Poisson distribution with an additional parameter allowing the variance to exceed the mean.

In the Hausman, Hall and Griliches (1984) negative binomial model it is assumed that $y_{it} \mid \gamma_{it} \sim \text{Poisson}(\gamma_{it})$ and $\gamma_{it} \mid \theta_i \sim \text{Gamma}(\lambda_{it}, 1/\theta_i)$, where θ_i is the dispersion parameter and $\ln \lambda_{it} = \beta \mathbf{x}_{it}$. This yields the following density function:

$$f(y_{it} \mid \lambda_{it}, \theta_i) = \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)} \left(\frac{1}{1+\theta_i} \right)^{\lambda_{it}} \left(\frac{\theta_i}{1+\theta_i} \right)^{y_{it}} \tag{18}$$

where Γ is the gamma function. Looking at the within-group effects only, this specification yields a negative binomial model for the *i*-th individual with

$$\begin{aligned} E(y_{it}) &= \theta_i \lambda_{it} \\ V(y_{it}) &= (1+\theta_i)\theta_i \lambda_{it} \end{aligned} \tag{19}$$

Under this model the ratio of the variance to the mean (dispersion) is constant within group and equal to $(1+\theta_i)$.

Hausman, Hall and Griliches (1984) further assume that for each individual i the y_{it} are independent over time. This implies that $\sum_t y_{it}$ also has a negative binomial distribution with parameters θ_i and $\sum_t \lambda_{it}$. Conditioning on the sum of counts, the resulting likelihood function for a single individual is

$$\frac{\Gamma\left(\sum_t y_{it} + 1\right)\Gamma\left(\sum_t \lambda_{it}\right)}{\Gamma\left(\sum_t y_{it} + \sum_t \lambda_{it}\right)} \prod_t \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)} \tag{20}$$

which is free of the θ_i parameters. The likelihood for the entire sample is then obtained by multiplying all the individual terms like (20) and can be maximised with respect to β the parameters using conventional numerical methods.

Unfortunately, this conditional negative binomial model is not a true fixed-effects method. In a recent paper, Allison and Waterman (2002) have proven that this method does not in fact control for all stable covariates. They argue that the problem originates from the fact that the θ_i parameters that are conditioned out of the likelihood function do not correspond to different intercepts in the log-linear decomposition of λ_{it}.

If we write $\theta_i = \exp(\delta_i)$, equations (19) imply that

$$E(y_{it}) = \exp(\delta_i + \beta \mathbf{x}_{it})$$
$$V(y_{it}) = (1 + e^{\delta_i}) E(y_{it})$$

from which it appears that the model does allow for an arbitrary intercept δ_i for each individual. However, while changes in $\mathbf{x}_{it}$ affect the mean directly and affect the variance only indirectly through the mean, changes in δ_i affect the variance both indirectly, through the mean, and directly. If δ_i is regarded as representing the effect of omitted explanatory variables, then there is no reason why such variables should have a different kind of effect from that of $\mathbf{x}_{it}$.

Alternatively, starting from (19) suppose that

$$\lambda_{it} = \exp(\delta_i + \beta \mathbf{x}_{it} + \gamma z_i)$$

where δ_i is an individual specific intercept and z_i is a vector of time-invariant covariates. Then conditioning on the total count for each individual does not eliminate δ_i or z_i from the likelihood function[28].

Allison and Waterman (2002) explore alternative methods to control for the δ_i's in the presence of overdispersion. Among the possibilities examined by the authors, a simulation study yields good results from applying the conditional fixed-effects Poisson estimator or, alternatively, an unconditional negative binomial regression estimator (that is assuming that y_{it} has a negative binomial distribution with mean μ_{it} and overdispersion parameter λ) with dummy variables to represent the fixed effects. They show that this last estimator has generally better sampling properties than the fixed effects Poisson estimator and it does not suffer from the incidental parameter bias in the coefficients. However, since it is accompanied by underestimates of the standard errors, these need to be adjusted upward. The downward bias in the standard error estimates can be easily and effectively corrected using a correction factor based on the deviance statistics, where the deviance is defined as

$$D = \sum_i \sum_t \{y_{it} \ln(y_{it}/\mu_{it})\} - (y_{it} + \lambda)\ln[(y_{it} + \lambda)/(\mu_{it} + \lambda)]$$

4.5.2 Empirical results from the entire sample

Table 4.6 reports the coefficients and standard errors from the estimation of the basic and extended specification for the entire sample (i.e. all industries and all countries)[29]. Columns labelled FE and RE report results from the fixed-effects and random effects estimation of the log-linear version of the model; OLS results are reported for comparison in columns one and six. Columns

[28] Symptomatic of this problem is that using statistical packages like Stata and Limdep, which implement (20), one can estimate regression models with both an intercept and time-invariant covariates, which is usually not possible with conditional fixed-effects models.
[29] To allow identification of the own R&D effect, all the models include a dummy variable that controls for those micro-sectors with very few patents that are assigned to industries with high R&D expenditures. This added variable (not reported in the table with estimation results) is found to be most effective in OLS estimation, but almost irrelevant in the other models used.

labelled CNB report estimates from the conditional fixed effects negative binomial model proposed by Hausman, Hall and Griliches (1984). Finally, columns labelled UNB report estimates from the unconditional fixed effects negative binomial estimator, with standard errors corrected using the deviance statistics as explained in the previous section[30].

In both the basic and extended specifications the two spillovers variables are always found significant. However, while the size of the international spillover indicator is fairly similar in the different regression models, this is not the case for the inter-sector national spillovers indicator. The difference across specifications suggests that this variable might be (negatively) correlated with the individual effects[31]. This is mainly due to the high serial correlation in the national spillovers variable coupled with its high variability across individuals and gives a further reason for fixed effects estimates to be preferred. Note, however, that if the true flow of national spillovers to a micro-sector is indeed constant in time, then fixed effects estimates might overemphasise the effect of the noise around this value.

Concerns about the ability of the conditional negative binomial estimation to effectively control for the individual effects are confirmed by the result on the coefficient of *ns* which, although positive and significant, is closer to the OLS estimate than to the fixed effect one. By contrast, the estimate from the unconditional negative binomial model is remarkably close to the result from fixed effects estimation on the log-linear model. On this basis, the log linear

---

[30] Estimates from the fixed effects Poisson and negative binomial regressions show evidence of overdispersion in the data (the ratio of the deviance to the degrees of freedom is well above one in all cases, whereas for a good fitting model they should be close to 1). Besides, Allison and Waterman (2002) show that the unconditional fixed effects negative binomial estimator is virtually always a better choice than the fixed effects Poisson estimator. For these reasons, estimates from this last regression models are not reported.

[31] Note that the random effects estimation of the same log-linear model delivers estimates close to fixed effects for all coefficients, but the coefficient of *ns* (note however that fixed effects and random effects estimates cannot be directly compared through the Hausman test, since random effects is not efficient in this case).

fixed effects and the unconditional negative binomial specifications are to be preferred[32].

The last five columns of Table 4.6 present estimation results for the extended specification. This includes interactions between the spillover indicators and the variable accounting for the incidence of self citations, which is used here as a proxy for firm level research experience in technology related areas. Coefficients are remarkably stable across regression models[33] and past research efforts appear to be more effective in increasing the elasticity of patents to international spillovers (a simple F test of equality between the two interaction coefficients strongly rejects the null hypothesis). This might be related to the fact that the indicator of international spillovers includes both intra- and inter-sector knowledge flows, while *ns* only accounts for inter-sector knowledge flows. Unfortunately, the data do not allow estimating precisely two separate effects (inter-industry vs. intra-industry) for international spillovers, as that would considerably increase the correlation among some of the explanatory variables.

These results show that international spillovers play an important role in explaining innovative productivity: in the preferred specifications, their coefficient is always positive and comparable to that of national spillovers and of own R&D. The estimation results also provide evidence of a positive effect of past research effort on the ability to understand and exploit external knowledge, that is of a significant role of absorptive capacity in increasing innovative productivity. Indeed, the estimated overall elasticity of patents to absorptive capacity from the fixed effects linear model, evaluated at the mean

[32] A potential critique to this approach is that it does not account for the endogeneity of R&D. However, this objection might be stronger when firm level data are used. Furthermore, recognising the endogeneity of R&D would call for instrumental variables methods and, in the present context, lagged values of the series are the only available instruments. The benefits of employing these instruments are highly uncertain, as Generalised Method of Moments methods do not perform well when the series are persistent.

[33] OLS coefficients are qualitatively comparable, although larger in absolute value.

of the variables, is equal to 0.16. Because the coefficients on the interaction terms are multiplied by λ and are always positive, if anything this result underestimates the true elasticity of patents to absorptive capacity. Note, however, that its effect is comparable to the effect of own R&D.

Table 4.6 Regression results for the entire sample from the linear and non-linear models

	OLS	FE	RE	CNB	UNB	OLS	FE	RE	CNB	UNB
Rd	0.18	0.18	0.27	0.20	0.18	0.18	0.18	0.26	0.20	0.18
	(.06)	(.02)	(.02)	(.01)	(.02)	(.06)	(.02)	(.02)	(.01)	(.02)
Ns	-0.03	0.31	-0.02	0.06	0.34	-0.04	0.27	-0.02	0.05	0.32
	(.01)	(.04)	(.01)	(.01)	(.04)	(.01)	(.04)	(.006)	(.01)	(.04)
Is	0.26	0.24	0.25	0.30	0.26	0.21	0.22	0.23	0.29	0.26
	(.12)	(.03)	(.03)	(.03)	(.03)	(.12)	(.03)	(.03)	(.03)	(.03)
*Ns*self*						0.13	0.02	0.03	0.03	0.02
						(.03)	(.01)	(.006)	(.01)	(.01)
*Is*self*						0.22	0.07	0.07	0.05	0.07
						(.04)	(.01)	(.008)	(.01)	(.01)
self						-1.60	-0.46	-0.52	-0.46	-0.47
						(.36)	(.09)	(.09)	(.12)	(.13)
time effect	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
country effects	yes	n.a.	yes	yes	yes	yes	n.a.	yes	yes	yes
industry effects	yes	n.a.	yes	yes	yes	yes	n.a.	yes	yes	yes
LnLik				-30296	-33368				-30217	-33278
Obs.					10680					

Note: Columns labelled OLS, FE (fixed effects) and RE (random effects) report estimates of the linear model, where the dependent variable is ln(patents+1). Finally, columns labelled CNB and UNB report estimates from the conditional and unconditional negative binomial models, respectively. Estimates from the unconditional negative binomial model are obtained adding dummy variables to represent the individual effects (not reported). Standard errors are in parentheses. OLS standard errors are robust to heteroskedasticity and correlation within panels: both these and RE standard errors might be biased downwards as they do not account for the correlation across individual effects. FE and UNB standard errors are instead reliable. The latter are corrected using the deviance statistics.

4.5.3 Empirical results at the industry level

In the regressions on the entire sample industry dummies are found significant: this provides a first coarse indication of the existence of relevant differences across industries. In order to gain a more complete understanding on this issue, Table 4.7 presents results from industry level regressions.

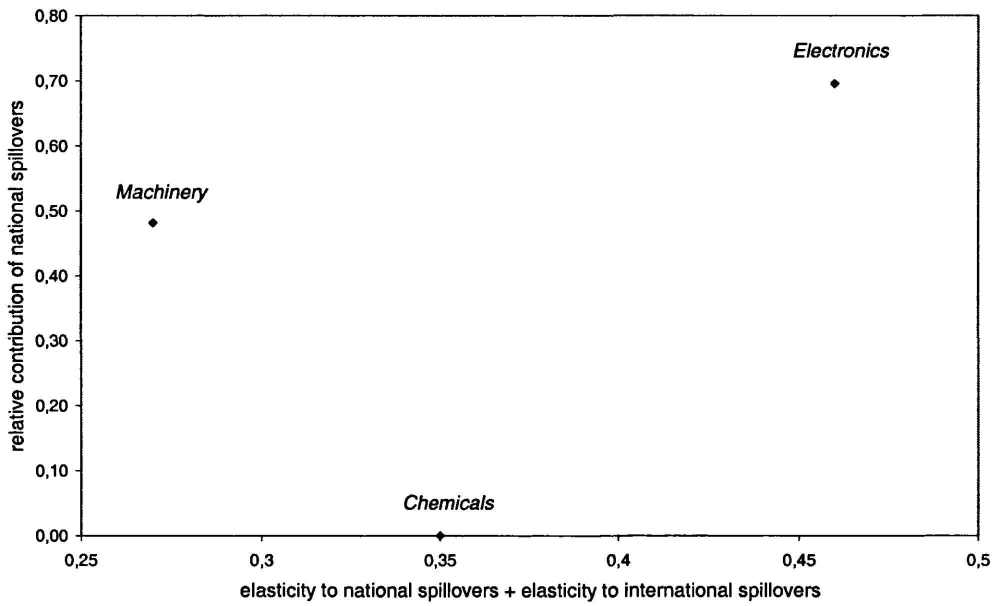
Table 4.7 Regression results at the industry level from the linear and non-linear models

	CHEMICALS			ELECTRONICS			MACHINERY		
	FE	**CNB**	**UNB**	**FE**	**CNB**	**UNB**	**FE**	**CNB**	**UNB**
rd	0.41	0.46	0.46	0.12	0.11	0.13	0.21	0.17	0.17
	(.05)	(.04)	(.06)	(.03)	(.02)	(.04)	(.03)	(.02)	(.03)
ns	0.06	-0.01	0.14	0.32	0.07	0.24	0.13	0.10	0.28
	(.07)	(.01)	(.09)	(.06)	(.01)	(.08)	(.05)	(.02)	(.05)
is	0.35	0.38	0.33	0.14	0.21	0.14	0.14	0.25	0.19
	(.07)	(.05)	(.06)	(.05)	(.04)	(.05)	(.06)	(.05)	(.05)
*ns*self*	0.04	0.04	0.03	-0.004	0.01	-0.002	0.006	0.02	-0.003
	(.01)	(.01)	(.01)	(.02)	(.02)	(.03)	(.01)	(.02)	(.02)
*is*self*	0.09	0.09	0.10	0.09	0.06	0.08	0.04	0.04	0.05
	(.01)	(.01)	(.01)	(.02)	(.02)	(.03)	(.02)	(.02)	(.02)
self	-0.82	-0.91	-0.93	-0.69	-0.59	-0.65	-0.21	-0.35	-0.27
	(.14)	(.17)	(.20)	(.21)	(.26)	(.33)	(.18)	(.27)	(.27)
time effect	yes	yes	yes	yes	yes	yes	yes	yes	yes
logLik		-11571	-12742		-9751	-10748		-8783	-9692
Obs.	4290	4290	4290	3270	3270	3270	3120	3120	3120

Note: See Table 4.6.

Micro-sectors within the chemical industry display a high elasticity to own R&D compared to micro-sectors in the electronics and machinery industries. Inter-sector national spillovers are never found effective in increasing innovation independently of absorptive capacity in the chemical industry, while in the electronics industry their impact is stronger than that of own R&D[34]. Finally, the elasticity of patents to international spillovers is always positive and significant and it is not statistically different from that to own R&D: a test of equality between the coefficients of *rd* and *is* cannot reject the null in each of the three samples.

Figure 4.1 Relative importance of national and international spillovers in the three industries



These estimates show that, with the exception of chemicals, national and international spillovers are together more effective than own R&D in increasing innovative performance. However, their relative importance is different in the three industries: international spillovers are respectively more, equally and less effective than national spillovers in the chemicals, machinery and electronics industry[35], as summarised in Figure 4.1.

---

[34] A test of equality between the coefficients of *rd* and *ns* rejects the null at the 5 percent confidence level.

[35] Although in the machinery sample the point estimate of the coefficient of *ns* from the

With reference to absorptive capacity, the results show that it is effective in rising the elasticity of patents to international spillovers in all industries. The overall elasticity of patents to absorptive capacity obtained from the estimated linear model and calculated around the means of the variables is equal to 0.27 in chemicals, 0.13 in electronics and 0.09 in machinery. Hence own past experience in technology related fields seems to be particularly important in the chemicals industry, where a unit increase in the indicator of experience would generate 45 more patents in the current year at the mean of the variables. This is almost the double of the average number of patents in the chemicals sample. A unit increase in the indicator of experience would instead generate 59 more patents in the electronic industry (1.3 times the average) and 29 more patents in the machinery industry (about the average number of patents in the industry).

4.5.4 Leaders vs. "followers"

Looking both at the volume of patent applications (Table 4.1) and at the direction of patent citations (Table 4.3) it is clear that US, Japan and Germany have the role of technological leaders and that France, the UK and Italy, although definitely among the most advanced countries, are somewhat lagging behind. Based on this observation, I split the sample in two groups, leaders (US, Japan and Germany) vs. "followers" (France, UK and Italy), and perform separate estimations on the two samples.

The main interest here lies in assessing whether absorptive capacity has a different effect in the two groups. From the theoretical point of view, absorptive capacity can be thought of having a non-linear effect. The further a

---

unconditional negative binomial model appears larger than the estimate of the coefficient of *is*, the difference is not statistically significant. Note also that the higher relative weight of national spillovers over international spillovers might help explaining the higher persistence found in electronics in chapter 2, althought the classification used there is different from the one employed here.

firm/country is from the technological frontier (i.e. the larger the gap with the technological leaders), the lower is its ability to absorb and exploit new external knowledge (mostly produced from the technological leaders). However, the farther a country is from the technological frontier, the larger is its potential to increase this ability (Griffith et al, 2000).

Table 4.8 Regression results for different groups of countries

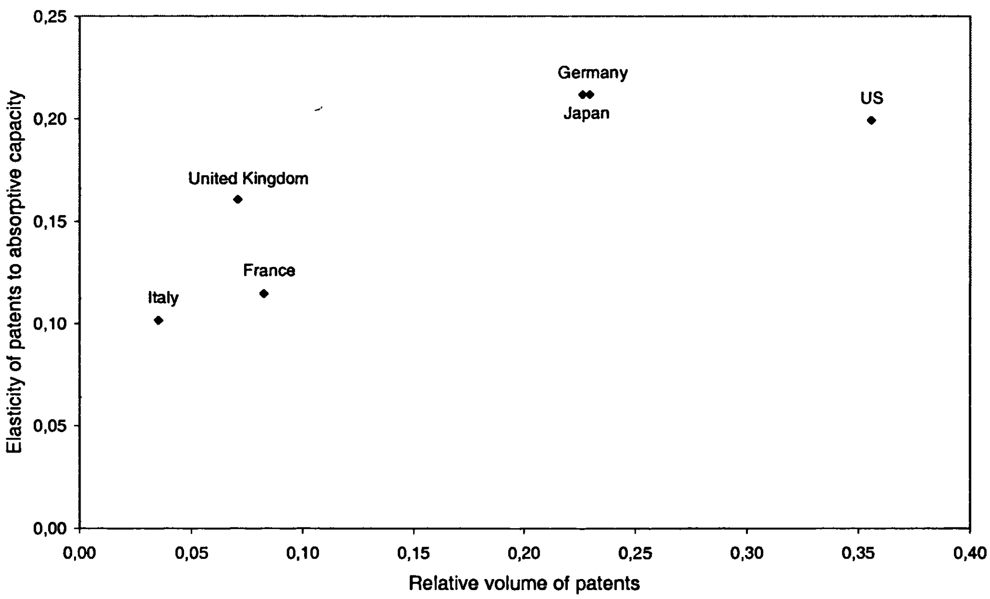
	LEADERS			"FOLLOWERS"		
	FE	CNB	UNB	FE	CNB	UNB
rd	0.24	0.24	0.23	0.17	0.13	0.13
	(.03)	(.02)	(.03)	(.03)	(.03)	(.03)
ns	0.24	0.06	0.30	0.15	0.04	0.21
	(.05)	(.01)	(.06)	(.05)	(.01)	(.05)
is	0.23	0.31	0.25	0.32	0.46	0.47
	(.04)	(.03)	(.03)	(.07)	(.06)	(.06)
*ns*self*	0.02	0.02	0.001	0.02	0.04	0.03
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
*is*self*	0.07	0.03	0.05	0.07	0.06	0.07
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
self	-0.29	-0.07	-0.03	-0.55	-0.72	-0.71
	(.12)	(.16)	(.22)	(.13)	(.17)	(.17)
time effect	yes	yes	yes	yes	yes	yes
country effects	n.a.	yes	yes	n.a.	yes	yes
industry effects	n.a.	yes	yes	n.a.	yes	yes
lnLik		-18121	-19934		-12064	-13320
Obs.	5685	5685	5685	4995	4995	4995

Note: See Table 4.6.

We would then expect to find a stronger overall elasticity of innovation (patents) to absorptive capacity in the group of technological leaders, compared to the "followers" (prediction 1). We would also expect the elasticity to absorptive capacity to increase less then proportionally as we move towards the technological frontier (prediction 2).

The estimation results for the two groups of countries are presented in Table 4.8. Technological leaders display elasticities to national and international spillovers similar to that of own R&D, while "followers" benefit more from international spillovers than from own research efforts (although the difference is significant only at the 10 percent confidence level).

Figure 4.2 Elasticity of patents to absorptive capacity



Note. The relative volume of patents is calculated with reference to the six countries total volume over the whole sample period.

In line with our expectations, the overall elasticity of patents to absorptive capacity is estimated to be 0.21 for leaders and 0.13 for "followers": a unit increase in the indicator of absorptive capacity at the means of the variables originates an increase in the number of patents equal to 76 in the technological leaders and to 17 in the "followers". In Figure 4.2, estimates of the elasticity of patents to absorptive capacity (calculated separately for each country in the sample) are plotted against the countries' relative volume of patents (a very coarse proxy for the world technological frontier). The resulting pattern appears increasing, thus in line with prediction 1. However the number of countries is too small to allow any clear inference on prediction 2, but note that the results are not inconsistent with the corresponding claim: the pattern also appears to increase at a declining rate, thus suggesting that a unit movement towards the technological frontier has a larger impact on the ability to absorb and exploit external knowledge the farther from the frontier itself is the country's initial position.

4.6 Conclusions

This chapter provides an empirical assessment of the effect of national and international knowledge spillovers on innovation at a finely defined sectoral level for six major industrialised countries over the period 1981-1995. Despite some data limitations, the results presented give evidence of the importance of such spillovers and of their different impact in different industries.

The measures of knowledge spillovers are built using citations included in patent applications at the European Patent Office. Once self-citations are controlled for, citation patterns do not show any home country bias. A large share of the total number of citations by patent applications from (firms within) a country are to foreign patents (international citations), the share being larger for countries behind the technological frontier. Consistently, international spillovers are always found to be effective in increasing innovative productivity.

The chapter then investigated the role of prior R&D experience in enhancing a country's ability to understand and improve upon external knowledge. This *absorptive capacity* is measured using self-citations, which are a signal of knowledge accumulation within the firm. The empirical results show that absorptive capacity increases the elasticity of a country's innovation to both national and international spillovers. Its effect is non-linear: the larger the gap of a country with the technological leaders the weaker is the country's ability to absorb and exploit external knowledge, but the larger is its potential to increase such ability.

Appendix

Table 4.9 R&D data aggregation from the OECD/ANBERD database.

ISIC Rev. 2	
31	Food, Beverages & Tobacco
32	Textiles, Apparel & Leather
33	Wood Products & Furniture
34	*Paper, Paper Products & Printing*
351+352-3522	Chemicals excl. Drugs
3522	Drugs & Medicines
353+354	Petroleum Refineries & Products
355+356	Rubber & Plastic Products
36	Non-Metallic Mineral Products
371	Iron & Steel
372	Non-Ferrous Metals
381	Metal Products
382-3825	Non-Electrical Machinery
3825	Office & Computing Machinery
3830-3832	Electric. Machin. excluding Commercial Equipment
3832	Radio, TV & Communication Equipment
3841	Shipbuilding & Repairing
3843	Motor vehicles
3845	Aircraft
[illegible]	[illegible]
39	Other Manufacturing

The 135 micro-sectors employed in the analysis belong (entirely or partially) to the sectors whose rows have been evidenced. In only one case (one electronics micro-sector in the UK) I have used R&D data for "Paper, Paper Products & Printing".

Table 4.10 Relative share of number of citations per patent within (intra-class) and outside (inter-class) the micro-sector of the citing patent.

Country(*)	Sector(*)	Intra-class	Inter-class
All	*All*	0.59	0.41
	Chemicals	0.56	0.44
	Electronics	0.64	0.36
	Machinery	0.56	0.44
Germany	*All*	0.58	0.42
	Chemicals	0.56	0.44
	Electronics	0.63	0.37
	Machinery	0.56	0.44
France	*All*	0.59	0.41
	Chemicals	0.55	0.45
	Electronics	0.64	0.36
	Machinery	0.56	0.44
Italy	*All*	0.60	0.40
	Chemicals	0.57	0.43
	Electronics	0.63	0.37
	Machinery	0.60	0.40
Japan	*All*	0.59	0.41
	Chemicals	0.55	0.45
	Electronics	0.62	0.38
	Machinery	0.53	0.47
UK	*All*	0.57	0.43
	Chemicals	0.54	0.46
	Electronics	0.63	0.37
	Machinery	0.56	0.44
US	*All*	0.61	0.39
	Chemicals	0.57	0.43
	Electronics	0.66	0.34
	Machinery	0.58	0.42

(*) Country and Sector refer to the citing patent.

REFERENCES

Aghion, P. and P. Howitt (1998), "Endogenous Growth Theory", Cambridge: The MIT Press.

Allison, P.D. and R.P. Waterman (2002), "Fixed-Effects Negative Binomial Regression Models", *Sociological Methodology* 32 (1): 247-265.

Amendola, G., G. Dosi and E. Papagni (1993), "The Dynamics of International Competitiveness", *Weltwirtschaftliches Archiv* 129 (3): 451-471.

Amendola, G., P. Guerrieri and P.C. Padoan. (1998), "International Patterns of Technological Accumulation and Trade" in D. Archibugi and Michie (eds.), "Trade, Growth and Technical Change", New York and Melbourne: Cambridge University Press.

Amiti, M. (1999), "Specialization Patterns in Europe", *Weltwirtschaftliches Archiv* 135 (4): 573-593.

Anderson, P. and M.L. Tushman (1990), "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change", *Administrative Science Quarterly* 35 (4): 604-633.

Anderson, T.W. and L.A. Goodman (1957), "Statistical Inference About Markov Chains", *Annual Mathematical Statistics* XXVIII.

Archibugi, D. and M. Pianta (1992a), "Specialization and Size of Technological Activities in Industrial Countries: The Analysis of Patent Data", *Research Policy* 21 (1): 79-93.

Archibugi, D. and M. Pianta (1992b), "The Technological Specialization of Advanced Countries", Norwell, Mass. and Dordrecht: Kluwer Academic; Brussels: Commission of European Communities.

Archibugi, D. and M. Pianta (1994), "Aggregate Convergence and Sectoral Specialization in Innovation", *Journal of Evolutionary Economics* 4 (1): 17-33.

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

Arthur, W.B. (1989), "Competing Technologies, Increasing Returns and Lock-in by Historical Events", *Economic Journal* 99 (394): 116-131.

Arthur, W.B. (1994) "Increasing Returns and Path Dependence in the Economy", The University of Michigan Press.

Balassa, B. (1965), "Trade Liberalization and Revealed Comparative Advantage", *The Manchester School of Economic and Social Studies* 33 (2): 99-123.

Basawa, I.V. and B.L.S. Prasaka Rao (1980), "Statistical Inference for Stochastic Processes", London: Academic Press.

Bernstein, J.I. and M.I. Nadiri (1989), "Research and Development and the Intra-industry Spillovers: An Empirical Application of Dynamic Duality", *Review of Economic Studies* 56 (2): 249-269.

Bottazzi, L. and G. Peri (2003), "Innovation and Spillovers in Regions: Evidence from European Patent Data", *European Economic Review* 47 (4): 687-710.

Branstetter, L.G. (2001), "Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the U.S. and Japan", *Journal of International Economics* 53 (1): 53-79.

Brasili, A., P. Epifani and R. Helg (2000), "On the Dynamics of Trade Patterns", *De-Economist* 148 (2): 233-57.

Breschi, S. and F. Lissoni (2001) "Knowledge Spillovers and Local Innovation Systems: A Critical Survey", *Industrial and Corporate Change* 10 (4): 975-1005

Breschi, S., F. Malerba and L. Orsenigo (2000), "Technological Regimes and Schumpeterian Patterns of Innovation", *Economic Journal* 110 (April): 388-410.

Bulli, S. (2001), "Distribution Dynamics and Cross-country Convergence: A New Approach", *Scottish Journal of Political Economy* 48 (2): 226-243.

Caballero, R. and A.B. Jaffe (1993), "How High are the Giant's Shoulders", NBER Working Paper no. 4370.

Cameron, A.C. and P.K. Trivedi (1998), "Regression Analysis of Count Data", Cambridge University Press, Cambridge, UK.

Cantwell, J. (1989), "Technological Innovation and Multinational Corporations". Oxford and Cambridge, Mass.: Blackwell.

Cefis, E. and L. Orsenigo (2001), "The Persistence of Innovative Activities: A cross-country and cross-sector Analysis", *Research Policy* 30 (7): 1139-1158.

Cincera, M. and B. Van Pottelsberghe de la Potterie (2001), "International R&D Spillovers: A Survey", *Cahiers Economiques de Bruxelles* 169: 3-32.

Coe, D.T. and Helpman E (1995), "International R & D Spillovers", *European Economic Review* 39 (5): 859-887

Cohen, W.M. and D.A. Levinthal (1990), "Absorptive Capacity: A New Perspective on Learning and Innovation", *Administrative Science Quarterly* 35 (1): 128-152.

Cowell, F.A. (1995), "Measuring Inequality", London, New York and Toronto: Simon and Schuster International, Harvester Wheatsheaf./Prentice Hall.

Crépon, B. and E. Duguet (1997), "Estimating the innovation function from patent numbers: GMM on count panel data", *Journal of Applied Econometrics* 12 (3): 243–263.

David, P.A. (1986), "Understanding the Economics of QWERTY: The Necessity of History", in W.N. Parker "Economic History ad the Modern Economist", Oxford: Blackwell.

De Benedictis, L. ans M. Tamberi (2001), "A Note on the Balassa Index of Revealed Comparative Advantage", Quaderni di Ricerca del Dipartimento di Economia n. 158, Università di Ancona.

Dosi, G. (1988), "Sources, Procedures and Microeconomic Effects of Innovation", *Journal of Economic Literature* 26 (3): 1120-71.

Dosi, G., K. Pavitt and L. Soete eds. (1990), "The Economics of Technical Change and International Trade", Harvester Wheatsheaf: Hertfordshire.

Eaton, J., R. Evenson, S. Kortum, P. Marino and J. Putnam (1999), "Technological Specialization in International Patenting", in G. Saxonhouse and T.N. Srinivasan (eds.), "Development, Duality and International Economic Regime: Essays in Honor of Gustav Ranis", Ann Arbor: Michigan University Press.

Eaton, J. and S. Kortum (1996); "Trade in Ideas: Patenting and productivity in the OECD", *Journal of International Economics* 40 (3-4): 251-278.

Eaton, J. and S. Kortum. (1999); "International Technology Diffusion: Theory and Measurement", *International Economic Review* 40 (3): 537-570.

Freeman, C. (1982), "The Economics of Industrial Innovation", London: Frances Printer.

Friedman, M. (1992), "Do Old Fallacies Ever Die?", *Journal of Economic Literature* 30 (4): 2129-2132.

Geweke, J., Marshall, R. and Zaekin, G. (1986), "Mobility Indices in Continuous Time Markov Chains", *Econometrica* 54 (6):1407-23.

Greenhalgh, C. (1990), "Innovation and Trade Performance in the UK", *Economic Journal* 100 (400): 105-118.

Greenhalgh, C., P. Taylor and R. Wilson (1994), "Innovation and Export Volumes and Prices – A Disaggregated Study", *Oxford Economic Papers* 46 (1): 102-134.

Greenhalgh, C., G. Mavrotas and R. Wilson (1996), "Intellectual Property, Technological Advantage and Trade Performance of UK Manufacturing Industries", *Applied Economics* 28 (5): 509-519.

Griffith, R., R. Harrison and J. Van Reenen (2003), "Technology Sourcing by UK Manufacturing Firms: An Empirical Analysis Using Firm-level Patent Data", unpublished manuscript.

Griffith, R., S. Redding and J. Van Reenen (2001), "Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries", CEPR Discussion Paper no. 2457.

Griliches, Z. (1979), "Issues in Assessing the Contributions of Research and Development to Productivity Growth", *Bell Journal of Economics* 10 (1): 92-116.

Griliches, Z. (1990), "Patent Statistics as Economic Indicators: a Survey", *Journal of Economic Literature* 28 (4): 1661-1707.

Griliches, Z. (1998), "R&D and Productivity: The Unfinished Business", in Z. Griliches "R&D and Productivity", The University of Chicago Press.

Grossman, G.M. and E. Helpman (1991), "Innovation and Growth in the Global Economy", Cambridge and London: The MIT Press.

Gruber, W., D. Mehta and R. Vernon (1967), "The R&D Factor in International Trade and International Investment of US Industries", *Journal of Political Economy* 75 (1): 20-37.

Gustavsson, P., P. Hansson and L. Lundberg (1999), "Technology, Resource Endowments and International Competitiveness", *European Economic Review* 43 (8): 1501-1530.

Hall, B. and J. Mairesse (1995), "Exploring the Relationship Between R&D and Productivity in French Manufacturing Firms", *Journal of Econometrics* 65 (1): 263-293.

Hall, B., Z. Griliches and J. Hausman (1986), "Patents and R&D: Is There a Lag?", *International Economic Review* 27 (2): 265-283.

Harrigan, J. (1997), "Technology, Factor Supplies and International Specialization: Estimating the Neoclassical Model", *American Economic Review* 87 (4): 475-494.

Hausman, J., B. Hall and G. Griliches (1984), "Econometric Models for Count Data and an Application to the Patents-R&D Relationship", *Econometrica* 52 (4): 909-938.

Hotelling, H. (1932), "Review of The Triumph of Mediocrity in Business by Horace Secrist", *Journal of the American Statistical Association* 28 (184): 463-465.

Jaffe, A.B. (1986), "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value", *American Economic Review* 76 (5): 984-1001.

Jaffe, A.B. (1989), "Real Effects of Academic Research", *American Economic Review* 79 (5): 957-70.

Jaffe, A.B. and M. Trajtenberg (2002), "Patents, Citations and Innovations", MIT Press, Cambridge, Ma.

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993), "Geographic Localisation of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics* 108 (3): 577-598.

Johnson, D.K.N. (2002), "The OECD Technology Concordance (OTC): Patents by Industry of Manufacture and Sector of Use", STI Working Paper Series 2002/5.

Keesing, D. (1967), "The Impact of Research and Development on United Stated Trade", *Journal Of Political Economy* 75 (1): 38-48.

Keller, W. (1998), "Are International R&D Spillovers Trade Related? Analyzing Spillovers Among Randomly Generated Trade Partners", *European Economic Review* 42 (8): 1469-1481.

Keller, W. (2000), "Geographic Localization of International Technology Diffusion", NBER Working Paper W7509.

Kortum, S. and J. Putnam (1997), "Assigning Patents to Industries: Tests of the Yale Technology Concordance", *Economic Systems Research* 9 (2): 161-175.

Krugman, P. (1979), "A Model of Innovation, Technology Transfer and the World Distribution of Income", *Journal Of Political Economy* 87 (2): 253-266.

Krugman, P. (1987), "The Narrow Moving Band, the Dutch Desease and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Scale Economies", *Journal of Development Economics* 27 (1-2): 41-55.

Krugman, P. (1991), "Increasing Returns and Economic Geography", *The Journal of Political Economy* 99 (3): 483-499.

Laursen, K. (2000), "Do Export and Technological Specialization Patterns Co-evolve in Terms of Convergence or Divergence? Evidence from 19 OECD Countries, 1971-1991", *Journal of Evolutionary Economics* 10 (4): 415-36.

Leontief W. (1953), "Domestic Production and Foreign Trade: The American Capital Position Re-examined", *Proceedings of the American Philosophical Society*, September: 332-349.

Lucas, R. E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics* 22 (1): 3-22.

Magnier, A. and J. Toujas-Bernate (1994), "Technology and Trade: Empirical Evidences from the Five Major Industrialized Countries", *Weltwirtschaftliches Archiv* 130 (3): 494-520.

Malerba, F. (1992), "Learning by Firms and Incremental Technical Change", *The Economic Journal* 102 (413): 845-859.

Malerba, F. and L. Orsenigo (1995), "Schumpeterian Patterns of Innovation", *Cambridge Journal of Economics* 19 (1): 47-65.

Malerba, F. and L. Orsenigo (1996), "Schumpeterian Patterns of Innovation are Technology Specific", *Research Policy* 25: 451-478.

Malerba, F., L. Orsenigo and P. Peretto (1997), "Persistence of Innovative Activities, Sectoral Patterns of Innovation and International Technological Specialization", *International Journal of Industrial Organization* 15 (6): 801-826.

Mansfield, E. (1985), "How Rapidly does New Industrial Technology Leak Out?", *Journal of Industrial Economics* 34 (2): 217-223.

Maurseth, P.B. and B. Verspagen (2002), "Knowledge Spillovers in Europe: A Patent Citation Analysis", *Scandinavian Journal of Economics* 104 (4): 531-45.

Moulton, B. R. (1986), "Random Effects and the Precision of Regression Estimates", *Journal of Econometrics* 32 (3): 385-97.

Munt, G. and H. Grupp (1995), "Koncordanz zwischen der internationalen Patent und Warenklassifikation", Fraunhofer-Institut fur Systemtechnik und Innovationsforschung Karlsruhe.

Nelson, R.R. (1993), "National Innovation Systems: A Comparative Analysis", Oxford; New York, Toronto and Melbourne: Oxford University Press.

Nelson, R.R. and S. Winter (1982), "An Evolutionary Theory of Economic Change", Cambridge (Mass.): The Belknap Press.

Pakes, A. and A. Griliches (1984), "Patents and R&D at the Firm Level: A First Look", in A. Griliches (ed.), "R&D, Patents and Productivity", The University of Chicago Press.

Patel, P. and K.L.R. Pavitt (1991), "Europe's Technological Performance", in Freeman, C., M. Sharp and Walker (eds.), "Technology and the Future of

Europe: Global Competition and the Environment in the 1990's", London: Frances Pinter Publishers.

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

Peri, G. (2003), "Knowledge Flows, R&D Spillovers and Innovation", unpublished manuscript.

Posner, M. V. (1961), "International Trade and Technical Change", *Oxford Economic Papers* 13 (3): 323-341.

Proudman, J. and S. Redding (1998), "Persistence and Mobility in International Trade", CEPR Discussion Paper 1802.

Proudman, J. and S. Redding (2000), "Evolving Patterns of International Trade", *Review of International Economics* 8 (3): 373-396.

Quah, D.T. (1993a), "Empirical Cross-Section Dynamics in Economic Growth", *European Economic Review* 37 (2-3): 426-434.

Quah, D.T. (1993b), "Galton's Fallacy and the Tests for the Convergence Hypothesis", *The Scandinavian Journal of Economics* 95 (4): 427-443.

Quah, D.T. (1996a), "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics", *Economic Journal* 106 (437): 1045-55.

Quah, D.T. (1996b), "Aggregate and Regional Disaggregate Fluctuations", *Empirical Economics* 21 (1): 137-59.

Quah, D.T. (1996c), "Convergence Empirics Across Economies with (some) Capital Mobility", *Journal of Economic Growth* 1 (1): 95-124.

Quah, D.T. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal of Economic Growth* 2 (1): 27-59.

Redding, S. (1999), "Dynamic Comparative Advantage and the Welfare Effects of Trade", *Oxford Economic Papers* 51 (1): 15-39.

Redding, S. (2002), "Specialization Dynamics", *Journal of International Economics* 58 (2): 299-334.

Romer, P. (1990), "Endogenous Technological Change", *Journal of Political Economy* 98 (5): 71-102.

Rosenberg, N. (1982), "Inside the Black Box", Cambridge: Cambridge University Press.

Schankerman, M. and A. Pakes (1986), "Estimates of the Value of Patent Rights in European Countries during the post-1950 Period", *Economic Journal* 96 (December): 1052-1076.

Schmookler, J. (1966), "Invention and Economic Growth", Cambridge, Mass.: Harvard University Press.

Schumpeter, J.A. (1912), "The Theory of Economic Development", Cambridge: Harvard Economic Studies.

Schumpeter, J.A. (1942), "Capitalism, Socialism and Democracy", New York: Harper and Brothers.

Shorrocks, A.F. (1978), "The Measurement of Mobility", *Econometrica* 46 (5): 1013-24.

Silverman, B. W. (1986), "Density Estimation for Statistics and Data Analysis", New York: Chapman and Hall.

Soete, L. (1981), "A General Test of Technological Gap Trade Theory", *Weltwirtschaftliches Archiv* 117 (4): 638-60.

Soete, L. (1987), "The Impact of Technological Innovation on International Trade Patterns: The Evidence Reconsidered", *Research Policy* 16 (2-4): 101-130.

Stern, R. and K. Maskus (1981), "Determinants of the Structure of US Foreign Trade 1958-76", *Journal of International Economics* 11 (2): 207-224.

Stolpe M. (1995), "Technology and the Dynamics of Specialization in Open Economies", Tuebigen: Mohr.

Vernon, V. (1966), "International Investment and International Trade in the Product Cycle", *Quarterly Journal of Economics* 80 (2): 190-207.

Weidlich, W. and G. Haag (1983), "Concepts and Models of a Quantitative Sociology: The Dynamics of Interacting Populations", Springer-Verlag.

Winkelman, R. (2000), "Econometric Analysis of Count Data", Springer