Diss. ETH No. 24165

The Relationship between Knowledge Absorption and Innovation Outcomes

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

presented by
Florian Seliger

M.A. Economics, M.Sc. Business Administration, LMU Munich

born on 25.05.1983

citizen of Germany

accepted on the recommendation of Prof. Dr. Peter Egger and PD Dr. Martin Wörter

2017

Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to thank Professor Peter Egger. His econometric expertise and innovative research ideas were irreplaceable in order to get started with the data collection and to deal with methodological issues. He is also an important contributor to one of the papers presented here. In deep gratitude, I want to mention Spyros Arvanitis and Martin Wörter from the KOF innovation group. Martin Wörter officially acted as co-supervisor, but Spyros Arvanitis was also involved in many steps and contributed with his knowledge in empirical innovation research. Both are co-authors of several papers. They also acted as great collaborators, colleagues, and persons. In the same vein, I want to mention Tobias Stucki, Thomas Bolli, and Benjamin Balsmeier, all of them co-authors and great colleagues. I also thank Andrin Spescha and Marius Ley for their support in data questions with respect to survey data, Mark Thompson and Christian Soltmann for their support with respect to the PATSTAT database and patent data in general, and Simon Bösenberg for the knowledge exchange with respect to patent data.

I want to thank the whole KOF Administration Team – and in particular Katharina Bloch – for taking care of all administrative tasks in a gentle way, Charles Clavadetscher for his support in databases, Tolga Coban and Daniel Hollenstein for general IT support. I am also grateful to discussants at conferences, especially Mark Lorenzen from DRUID for reading and discussing my papers. I also benefited from the support other researches offer on Statalist.

Finally, I want to thank Pauliina Sandquist, Andreas Dibiasi, Marius Ley and Kushtrim Veseli for their good company also outside KOF and the whole team from the KOF Education Systems group (Johanna Kemper, Maria-Esther Egg, Jutta Bürgi, Katherine Caves, Thomas Bolli and all interns and civil servants I cannot mention here) for their nice company during lunch time.

Last but not least, I want to mention my parents, my brother and old friends (Ingrid, Berthold, Christian, Philipp, Stefanie, Martin, Parichaya, Tobias, Miriam, Konrad).

1 Introduction

1.1 Motivation

1.1.1 External Knowledge Absorption, Knowledge Spillovers, and Open Innovation

The thesis deals with the internal and external knowledge absorption and innovation outcomes at two levels: the firm and the sector-country. Firms and other economic actors do not act in a vacuum, they collaborate and interact in various ways with each other. In this way, they gather knowledge from external sources in order to boost their innovation activities. Major channels include hiring talents from competitors or universities, external R&D contracts, licensing, collaboration agreements etc. (see, e.g., Aschhoff and Schmidt 2008; Belderbos et al. 2004; Cassiman and Veugelers 2002, 2006; Dhont-Peltrault and Pfister 2011; Tether 2002).
In addition to the more formalized ways of acquiring external knowledge, firms can also draw on knowledge spillovers in order to generate technological opportunities.[1] These are knowledge externalities that are usually not based on formal agreements but rather on informal and infrequent exchange of ideas or at least the exposition to an environment where ideas are created (e.g., if a firm is located in a so-called knowledge cluster where it is exposed to a pool of researchers, highly-qualified employees etc. that are active in this cluster and produce relevant knowledge). They are non-pecuniary and available in the knowledge environment for any actor who is able to absorb this kind of knowledge. Costs are imposed by the dissipation of own knowledge to competitors and the up-front investments in absorptive capacity that is needed to understand external knowledge and to generate new knowledge (see Cohen and Levinthal 1989, 1990). The strength of inter-firm spillovers usually depends on a firm's position in the technological and geographical space relative to other firms.

[1] see Griliches (1995); Griliches (1992)

Whereas the concept of knowledge spillovers has been mainly studied by economists[2], management scholars usually refer to the so-called Open Innovation paradigm. They argue that innovative firms draw on knowledge generated by a wide range of external sources such as customers, competitors, suppliers and universities (Chesbrough 2006; Laursen and Salter 2006; West and Bogers 2014; West et al. 2014). The idea is that a firm actively and strategically searches for external knowledge which contrasts with the spillover concept where a firm is merely exposed to an external pool of knowledge. In the latter case, research usually does not make any assumptions on whether a firm has an explicit search strategy or not and a firm might even benefit from spillovers involuntarily – at least by conception. However, the degree to which a firm can benefit from an external pool of knowledge depends on economic or technological interactions of the firm with the environment and the respective closeness to other actors. Therefore, the pool of knowledge is weighted appropriately in empirical investigations in order to reflect the probability of interactions in trade or knowledge exchange.

In my thesis, I mainly follow the economic path and look at the impact of knowledge spillovers on innovation success (Chapter 2) and the generation of further knowledge flows (Chapter 4). In Chapter 3, we look at the impact of diversity of a firm's technological portfolio on the number of patents and innovation success. Although diversity refers to the existing knowledge stock of a firm and not to external spillovers, it is relevant in this context as it can increase intrafirm spillovers and absorptive capacity that is needed in order to absorb external spillovers.

Innovation success is the main outcome that is studied, i.e. the commercial success of innovations on the market. Whenever possible, I differentiate between radical and incremental innovation success. Radical innovation – in contrast to incremental innovation – depends on the introduction of new concepts that depart significantly from prior and current inventions, may use a new core technology and has the potential to generate new markets and to influence future inventions (Chandy and Tellis 2000; Dahlin and Behrens 2005; Garcia and Calantone 2002).

1.1.2 Moderator Variables

In my thesis, I will look at different 'moderator variables' that might influence the relationships between some forms of knowledge absorption and innovation performance.

[2] e.g., Jaffe (1986). They also play an important role in endogenous growth theory (see Romer 1990).

Competition

The direct effect of competition on innovation has been studied extensively. One of the first empirical analysis of the effect of competition on innovation can be found – among others – in Scherer (1965). Aghion et al. (2005) offers a theoretical model in order to understand the inverted U-shaped relationship that has been found empirically (e.g., by Scherer 1967). Laggard firms catch-up with technologically leading firms in order to escape competition. Of course, catching up is only possible if there is some form of spillover from the leading to the laggard firms so that they can learn how to imitate the leaders' technology. Bloom et al. (2013) consider spillovers from product market in addition to spillovers from R&D activities and offer a stylized model for the 'business-stealing' effect that firms that benefit from knowledge exchange through knowledge spillovers face under product market rivalry because knowledge dissipates to competitors. However, there is still a lack of theoretical frameworks and empirical investigations taking into account the potentially complex relationships between knowledge spillovers and product market competition and their subsequent impact on innovation. Chapter 2 is an attempt in this direction and studies the effect of product market competition on the relationship between spillovers and innovation success.

Technological Uncertainty

Uncertainty about potential outcomes and other actors' actions is a central parameter in economic models. Surprisingly, in innovation economics uncertainty about technological discontinuities and competitors' technological advancements is rarely studied empirically. The literature on product innovation management, however, has incorporated this parameter in empirical investigations. We study the impact of technological uncertainty that a firm faces in its environment on the relationship between technological diversity and innovation in Chapter 3.

1.1.3 Goals

In my thesis, I want to make a couple of contributions to existing literature, both conceptually and empirically:

1. **Measures**

   In my thesis, I use well-known measures from literature for knowledge spillovers and technological distance. However, they are essentially adapted resp. modified (see Chapter 2 and Chapter 4 for details).

2. **Dependent Variables**

   Survey data provides us with a direct measure of innovation success – the sales share with innovative products – that is preferable to a count of patents that is often used in comparable studies. In addition, it is possible to differentiate between more incremental and more radical innovation success rather easily as this distinction is included in the respective survey questions (the sales share with innovations that are new to the firm as proxy for success with incremental innovations vs. the sales share with innovations that are new to the market as proxy for success with radical innovations).

3. **Moderator variables**

   In Chapter 2, we study the moderating effect of competition on the relationship between spillovers and innovation. In Chapter 3, we look at the interaction effect of technological diversity with technological uncertainty.

4. **Data**

   The Swiss Innovation Survey is one of the main data sources for this thesis. I matched firms from the survey with their patent applications and used the comprehensive information in patent citations, technological fields etc. In addition, I created a dataset at sector-country level that is used in Chapter 4 where patent citations are used in order to trace knowledge flows between input and output sector-countries.

1.2 Data and Methodology

My thesis uses data from mainly two sources: Survey data from the KOF Innovation Panel and patent data from PATSTAT, aggregated at firm, industry and country level. PATSTAT contains patent data from more than 100 patent offices over a time span of up to a century and allows for family and global patent analysis (De Rassenfosse et al. 2014). In Chapter 4 and 5, I use a couple of further data sources that are described in the respective sections. Applicants from PATSTAT with a Swiss country code were matched with firm names from the Innovation Panel.

I also identified firms in PATSTAT that were cited by Swiss firms in their patent applications (Chapter 2). For Chapters 4 and 5, PATSTAT data was consolidated at the country-industry level and matched with other country-industry data, e.g., from OECD.

More details on the data generation and the econometric methods that have been applied are described in detail in the respective Chapters.

1.3 Findings

1.3.1 Introductory Result

Table 1.3.1 reports partial correlations from a fractional logit model where we analyze the relationship between different modes of knowledge acquisition and two innovation outcomes: First, the sales share with products that are new to the firm and, second, the sales share with products that are new to the market. As already mentioned above, the former captures mainly incremental innovation as resulting from imitation. The latter captures more radical innovations (see Kleinknecht et al. 2002).[3]

Data is drawn from three cross-sections of the Swiss Innovation Panel (2005, 2008, 2011).[4] The variables capturing external knowledge sourcing closely follow Laursen and Salter (2006). Laursen and Salter examined the association of 'breadth' (number of knowledge sources that a firm uses) and 'depth' (intensity of use of those sources) of search for external knowledge with radical vs. incremental innovation success using data from the UK innovation survey. They found that 'breadth' shows a stronger (but curvilinear) association with incremental innovation success, whereas 'depth' is more important for radical innovation.

[3]The results are drawn from a KOF Working Paper (Seliger and Arvanitis 2014). The regressors are also described in detail there. In short: *rnd* – Incidence of R&D activities, *lnempl_shr_higher* – Share of employees with higher education, *lnempl_shr_train* – Share of employees with vocational training, *skill_imped* – Incidence of a lack of skilled employees, *tech_pot* – Available technological potential, *know_breadth* – Number of knowledge sources used, *know_breadth2* – Number of knowledge sources used, squared, *know_depth* – Number of important or very important knowledge sources, *know_depth2* – Number of important or very important knowledge sources, squared, *rnd_coop* – Incidence of R&D cooperation, *rnd_ext* – Incidence of external R&D, *demand* – Assessment of demand development, *price* – Strong or very strong price competition, *nprice* – Strong or very strong non-price competition, *n_compet_5* – Number of main competitors less than 5, *n_compet_15* – Number of main competitors between 6 and 15, *n_compet_50* – Number of main competitors between 16 and 50, *copy_imped* – Easiness of imitation, *fin_imped* – Lack of funding, *lninvest_pc* – Gross investments per employee, *lnict_inc_share* – Share of ICT investments, *lnempl* – Number of employees, *lnage* – Firm age, *foreign_owned* – Firm owned by foreign company. Similar to Egger and Kesina (2013, 2014), we display some criteria to evaluate the estimation quality of the model and the functional form of the model (correlation of actual and predicted outcomes, Akaike criterion, link test), for details see the references mentioned before.

[4]For details, see the KOF Working Paper.

Table 1.3.1: Marginal Effects From Fractional Logit Models, Innovating Firms

	inno_ share_new_firm	inno_ share_new_market
rnd	0.000	0.041***
	(0.01)	(0.01)
lnempl_shr_higher	-0.003	0.008
	(0.01)	(0.01)
lnempl_shr_train	0.011*	0.004
	(0.01)	(0.01)
skill_imped	0.011	-0.012
	(0.01)	(0.01)
fin_imped	-0.004	-0.004
	(0.01)	(0.01)
lninvest_pc	-0.001	-0.004
	(0.00)	(0.00)
lnict_inv_share	0.012**	0.007*
	(0.00)	(0.00)
demand	0.008	0.001
	(0.01)	(0.01)
price	0.004	0.001
	(0.01)	(0.01)
nprice	0.003	0.013*
	(0.01)	(0.01)
n_compet_5	-0.013	0.004
	(0.01)	(0.01)
n_compet_6_15	0.002	-0.004
	(0.01)	(0.01)
n_compet_16_50	0.013	-0.005
	(0.02)	(0.02)
copy_imped	-0.003	-0.010
	(0.01)	(0.01)
tech_pot	0.026***	0.008
	(0.01)	(0.01)
rnd_coop	0.019**	-0.002
	(0.01)	(0.01)
rnd_ext	0.017**	0.005
	(0.01)	(0.01)
know_breadth	0.012*	0.006
	(0.01)	(0.01)
know_breadth2	-0.001**	0.000
	(0.00)	(0.00)
know_depth	0.007*	0.006
	(0.00)	(0.00)
know_depth2	0.000	0.000
	(0.00)	(0.00)
lnempl	-0.004	0.006*
	(0.00)	(0.00)
lnage	-0.004	-0.016***
	(0.00)	(0.01)
foreign_owned	0.020**	-0.016*
	(0.01)	(0.01)
industry dummies	yes	yes
year dummies	yes	yes
N	1906	1906
AIC	0.681	0.648
chi2 (p-value)	0.000	0.000
Correlation real and predicted share of sales	0.300	0.301
Link test		
Sales share hat	1.557	1.298
	(0.921)	(0.997)
Sales share hat squared	0.185	0.098
	(0.274)	(0.263)

*** p<0.01, ** p<0.05, * p<0.1
cluster robust standard errors in parentheses

Our findings for Switzerland indicate that firms with incremental innovations are significantly more 'extroverted' than other innovating firms. Their innovation success as measured by the sales share of products 'new to the firm' is not only more related to the 'breadth' of external sourcing, but also to external R&D activities, R&D cooperation and medium-educated personnel as measured by the share of employees with vocational training. Radically innovating firms do not rely to the same extent on the exploration of external knowledge. Their rather 'introverted' behavior seems to be more related with intense exploitation of internal resources based on in-house R&D and personnel with tertiary-level education. This finding suggests that 'radical' innovations are not predominantly created by searching for external knowledge but by an innovator's successful commercialization of a single unique idea (see Cappelli et al. 2014; Garriga et al. 2013; Köhler et al. 2012).

1.3.2 Main Findings

In sum, the results depicted above are supported by the evidence found in the other papers that are presented in this dissertation. The results suggest that knowledge absorption (through formal or informal acquisition, through 'voluntary' or 'involuntary' knowledge absorption) is much more isolated than often suggested. For example, in Chapter 2, we find evidence that more innovative firms rely to a larger extent on internal knowledge stocks compared to firms with more incremental innovations. For the latter, spillovers from an external pool of knowledge are relatively more important. Chapter 2 also shows that the magnitude of the effect of spillovers depends on prior technological connections (we capture them with patent backward citation links) to potential spillover sources. Spillovers are much more effective when they come from 'relevant' or 'familiar' sources (in contrast to spillovers that are associated with the whole environment a firm is exposed to in a country or industry, which makes the spillover source more or less arbitrary).

In the same vein, Chapter 4 shows that sector-countries that are connected through an 'input-output' knowledge relationship will produce more knowledge flows if they have already exchanged knowledge so that knowledge is already 'familiar'. In contrast, rather few sector-country constellations can benefit from external knowledge spillovers, i.e. from spillovers from sector-countries that are not part of an existing relationship. These are mainly sector-countries where the generation of knowledge flows needs more advanced knowledge than provided in the respective constellation so that a kind of 'learning' from more advanced sources must take place. Chapter 3 shows that technological diversity (which may act as enabler for the absorption of external knowledge as well as for intrafirm spillovers) leads to more patented inventions, but to a lower share of sales with innovations. In the light of the other results, investing in multiple technological fields creates a double-edge sword as such investments might enhance absorptive capacity but do not contribute to success on the market with the underlying inventions in the short run. This Chapter also points to the importance of coordination costs within a firm that might play a role with respect to intrafirm spillovers. Clearly, there are also information costs involved that might prevent firms or countries from external knowledge absorption.

In sum, the results suggest that firms should aim at developing existing knowledge stocks in order to increase their innovative and absorptive capacity. They should be open to external knowledge sources, but absorption should take place in more familiar areas (i.e., technological areas that have been already exploited by the respective firm or by competitors that are generally active in similar areas) rather than increasing knowledge acquisition from multiple external actors or investing in very diverse areas at any cost.

1.4 Abstract of Scientific Papers

The thesis is a collection of papers, I wrote in collaboration with other researchers. They all revolve around the broad topic of the relationship between knowledge absorption and innovation outcomes. In what follows, each paper constitutes a chapter of the thesis.

The first paper (co-authored with Spyros Arvanitis and Martin Wörter) looks at the relationship between knowledge spillovers and innovation success. We propose a new patent-based measure of knowledge spillovers that calculates technological proximity based on firms that were identified via patent backward citation links. We argue that this measure has a couple of advantages as compared to the 'standard' measure proposed by Jaffe: First, it reflects spillovers from both domestic and foreign technologically 'relevant' firms, second, it is more precise because it only takes into account knowledge relations with technologically 'relevant' firms. Our empirical results indeed show that the measure performs better than the standard measure in an innovation model. We find – for a representative sample of Swiss firms – that knowledge spillovers measured in this way have a positive and significant impact on innovation success. However, the knowledge spillovers appear to be localized: Spillovers from geographically distant areas such as the USA and Japan matter less than spillovers from near destinations such as Europe and particularly Switzerland itself. Moreover, the spillover effect on innovation performance decreases with increasing number of competitors on the main product market so that this effect would appear only in niche markets or oligopolistic market structures. However, an additional effect of competition can only be detected for more radical innovation success.

The second paper (co-authored with Thomas Bolli and Martin Wörter) analyzes the impact of technological diversity in the patent portfolio on innovation performance in the discovery stage, measured by R&D intensity and patent applications, and in the commercialization stage, measured by the sales share generated by innovations, using Swiss firm-level panel data. While we do not find any impact of diversity on R&D intensity, we confirm a positive impact of diversity on patent applications as suggested by the literature. However, extending the analysis to the commercialization stage reveals that technological diversity decreases the share of sales generated by innovative products. Hence, technologically more diversified firms patent more but generate a smaller share of sales with new products. We argue that this pattern emerges because coordination costs induced by technological diversity matter more in the commercialization stage. Since technological uncertainty further increases coordination costs (in both the discovery and commercialization stage), this interpretation is in line with our additional result that high technological uncertainty increases the costs of diversity and negatively moderates the effect of diversity on innovation performance.

The third paper deals with spillovers and knowledge flows in a sectoral-country context. It studies determinants of knowledge flows as measured with patent forward citations that occur between 'input' and 'output sector-countries'. We look at the impact of absorptive capacity of a focal sector-country, knowledge spillovers and technological distance between sector-countries on further knowledge flows. For this purpose, we develop a knowledge flow matrix similar to input-output tables in trade where patent citations capture knowledge flows that go from the input sector-country to the output sector-country. We estimate a gravity model with variables that capture technological distance and knowledge that comes from either inside the input output pair or from external spillover sources. Our results indicate that knowledge originating from the output sector-country and - in some cases - external spillovers are key in generating further knowledge flows that go to the output sector-country. A distinction between high-tech and low-tech sector-countries shows that spillovers are more useful for the generation of knowledge flows if the input sector-country is low-tech. Low-tech sector-countries benefit from both high-tech knowledge from the output sector-country and external knowledge from the technological frontier. In contrast, knowledge flows based on high-tech sector-countries cannot benefit from low-tech sector-countries and only to a very limited extent from other high-tech sources. Technological distance between sector-countries has a negative impact on further knowledge flows so that only technologically proximate sector-countries are more likely to generate knowledge flows.

Finally, the fourth paper (co-authored with Peter Egger and Martin Wörter) is a more methodological paper that applies advanced estimation techniques on patent citation distributions. It analyzes features of the distribution of 3.7 mn. patent family applications from PATSTAT in 34 countries, 17 industries, and 11 years. Power-law regressions suggest that higher levels of R&D intensity and outward foreign direct investment in a country and a sector reduce the concentration of patent citations.

1.5 Zusammenfassung der Arbeitspapiere

Die Dissertation ist eine Zusammenstellung verschiedener Arbeitspapiere, die ich mit verschiedenen Autoren zusammen geschrieben habe. Sie kreisen alle um das breite Thema des Zusammenhangs zwischen Wissensabsorption und Innovationsergebnissen. Im Folgenden wird jedes Papier als ein Kapitel innerhalb der Dissertation dargestellt.

Das erste Papier dieser Dissertation (verfasst zusammen mit Spyros Arvanitis und Martin Wörter) behandelt die Beziehung zwischen Wissens-Spillover[5] und Innovationserfolg. Wir schlagen ein neues Mass für Wissens-Spillover vor, das auf Patentdaten basiert und die technologische Nähe zu Firmen, die in sogenannten Rückwärtszitationen in Patentanmeldungen aufscheinen, einbezieht. Wir argumentieren, dass dieses Mass eine Reihe von Vorteilen besitzt im Vergleich zu Standardmassen aus der Literatur. Erstens spiegelt es sowohl Spillovers von heimischen als auch ausländischen Firmen wider. Zweitens ist es präziser, da es ausschliesslich Wissensbeziehungen mit technologisch relevanten Firmen einbezieht. Unsere empirischen Resultate zeigen, dass dieses Mass in einem Innovationsmodell tatsächlich besser zu funktionieren scheint als das Standardmass. Wir finden – für eine repräsentative Stichprobe Schweizer Firmen –, dass

[5] Wissens-Spillovers sind externe Effekte, die entstehen, wenn ökonomische Akteure Wissen produzieren, von dem andere kostenlos profitieren können. Sie werden in Abschnitt 1.1.1 kurz erläutert. Ich behalte in der deutschen Zusammenfassung den englischen Begriff bei, um Mehrdeutigkeiten der Begrifflichkeit auszuschliessen.

Wissens-Spillover einen positiven und signifikanten Effekt auf den Innovationserfolg haben. Die Wissens-Spillover scheinen jedoch lokal begrenzt zu sein: Spillovers, die von geographisch weit entfernten Ländern kommen (z.B. aus den USA oder Japan), spielen eine geringere Rolle als Spillover von geographisch näheren Ländern in Europa. Darüber hinaus nimmt der Spillover-Effekt mit einer wachsenden Zahl von Wettbewerbern auf dem Hauptabsatzmarkt ab. Der Effekt sollte also nur für Nischenmärkte oder oligopolistische Marktstrukturen bestehen. Der Wettbewerbseffekt ist jedoch nur für den Innovationserfolg mit radikaleren Innovationen feststellbar.

Das zweite Papier (verfasst zusammen mit Thomas Bolli und Martin Wörter) analysiert den Einfluss von technologischer Diversität im Patentportfolio auf die Innovationsleistung in der Entdeckungsphase – gemessen durch F&E-Intensität und Patentanmeldungen – und in der Kommerzialisierungsphase – gemessen durch den Umsatzanteil mit Innovationen. Wir verwenden Schweizer Firmenpaneldaten. Während wir keinen Einfluss von Diversität auf die F&E-Intensität feststellen können, können wir – wie andere Autoren auch – einen positiven Einfluss von Diversität auf Patentanmeldungen bestätigen. Wenn wir die Analyse allerdings auf die Kommerzialisierungsphase ausweiten, sehen wir, dass technologische Diversität den Umsatzanteil mit innovativen Produkten senkt. Folglich patentieren diversifiziertere Firmen mehr, generieren jedoch einen kleineren Umsatzanteil mit neuen Produkten. Wir argumentieren, dass dieses Muster aufscheint, da Koordinationskosten, die durch technologische Diversität entstehen, in der Kommerzialisierungsphase grösseres Gewicht haben. Da technologische Unsicherheit die Koordinationskosten weiter erhöht (sowohl in der Entdeckungs- als auch in der Kommerzialisierungsphase), ist diese Interpretation vereinbar mit unserem zusätzlichen Ergebnis, dass hohe technologische Unsicherheit einen negativen Effekt auf die Beziehung zwischen Diversität und Innovationsleistung hat.

Das dritte Papier behandelt Spillovers und Wissensflüsse auf Länder-Sektoren-Ebene. Es untersucht Determinanten von Wissensflüssen zwischen Input- und Output-Ländersektoren. Wir betrachten den Einfluss der Absorptionsfähigkeit eines Ländersektors sowie von Wissens-Spillover und technologischer Distanz auf weitere Wissensflüsse zwischen den Ländersektoren. Hierfür entwickeln wir eine Wissensflussmatrix – ähnlich einer Input-Output-Tabelle im internationalen Handel. Patentzitationen messen hierbei Wissensflüsse von einem zum anderen Ländersektor. Wir schätzen ein Gravitätsmodell mit Variablen, die technologische Distanz und Wissen von innerhalb oder ausserhalb des Input-Output-Paares messen. Unsere Ergebnisse zeigen, dass Wissen, das ursprünglich dem Output-Ländersektor entstammt, und in manchen Fällen auch externe Spillovers entscheidend sind, um weitere Wissensflüsse hin zum Output-Ländersektor zu generieren. Eine Unterscheidung zwischen Hightech- und Lowtech-Ländersektoren zeigt, dass Spillovers nützlicher sind für die Generierung von Wissensflüssen, wenn der Input-Ländersektor ein Lowtech-Sektor ist. Lowtech-Ländersektoren profitieren sowohl von Hightech-Wissen als auch externem Wissen von der technologischen Front. Im Gegensatz dazu können Wissensflüsse, die von Hightech-Sektorländern ausgehen, nicht von Lowtech-Ländersektoren und nur zu einem geringen Grad von anderen Hightech-Quellen profitieren. Die technologische Distanz zwischen Ländersektoren hat einen negativen Einfluss auf weitere Wissensflüsse, so dass umgekehrt technologische Nähe die Wahrscheinlichkeit für Wissensflüsse erhöhen sollte.

Das letzte Papier, das ich zusammen mit Peter Egger und Martin Wörter verfasst habe, ist ein eher methodischer Beitrag, der fortschrittliche ökonometrische Schätzmethoden auf Verteilungen von Patentzitationen anwendet. Wir analysieren Eigenschaften der Verteilung von 3.7 Millionen Patentfamilienanmeldungen aus PATSTAT in 34 Ländern, 17 Industrien und 11 Jahren. Potenzgesetzschätzungen deuten darauf hin, dass höhere Niveaus der F&E-Intensität und ausländischer Direktinvestitionen in einem Land und Sektor die Konzentration der Patentzitationen senken.

2 Knowledge Spillovers and their Impact on Innovation Success - A New Approach Using Patent Backward Citations *

2.1 Introduction

Since the two seminal papers by Jaffe (1986) and Jaffe et al. (1993) patent-based measures of knowledge spillovers have become the workhorse in micro-level studies. Although Bloom et al. (2013) substantially extended the original Jaffe measure and made an effort to include spillovers from product market, the original approach to measure knowledge spillovers as suggested by Jaffe has sustained its attractiveness. With the paper at hand we suggest a modified version of the Jaffe measure and show its qualities in the framework of a standard innovation model. Moreover, we show that competition has a significant impact on the effect of spillovers on innovation performance.

In this paper, we also use the Jaffe approach to measure technological proximity between firms with the uncentered correlations between their underlying technological portfolios. A firm's technological portfolio is proxied by a firm's share of patent applications in technological fields according to the International Patent Classification (IPC). Annual patent flows are accumulated to patent stocks. The firms' patent stocks are then weighted with the patent-based Jaffe measure of technological proximity. For a focal firm the sum of these weighted patent stocks of the firms of the focal firm's technologically relevant environment is used as a proxy of potential knowledge spillovers in our innovation model.

However, we argue that the traditional Jaffe measure has two important drawbacks. First, it

*This chapter is co-authored with Spyros Arvanitis & Martin Wörter focuses on spillovers coming from firms belonging to a given sample. In most cases, such a sample is arbitrary and not representative of any relevant firm population. Furthermore, in most studies it is not possible to include spillovers from foreign firms although many patenting firms are acting globally and might benefit from knowledge generated elsewhere. Second, the traditional measure considers potential knowledge interactions with every firm in the sample, thus adding noise to the measure, even if many of these firms might not be technologically relevant for the focal firm.

Today's data availability and data processing capacities makes it possible to include much more firms that might be directly technologically relevant for a focal firm. For this exercise, the technological landscape of Switzerland is an ideal subject because Switzerland is a small country with a strongly internationalized economy. Therefore, we use a sample of firms with patent activities from the KOF Swiss Innovation Survey and search for links to other firms worldwide that are technologically relevant for the sample firms. Such technological links can be mapped with a focal firm's backward citations to another firm's patents. Thus, such backward links are used to identify technologically relevant firms worldwide. This is accomplished with the new names table in PATSTAT (the so-called ECOOM-EUROSTAT-EPO PATSTAT Person Augmented Table, EEE-PPAT table henceforth) that allows to identify cited firms. We then matched all patent applications that we found in PATSTAT with the cited firms, calculated their patent stocks and their patent shares in the underlying IPC classes.

We built N subsamples where N is the number of Swiss firms in our sample. Each subsample contains $1 + n_i$ firms where n_i is the number of firms cited by Swiss firm i. Based on this subsample we calculated the Jaffe measure for each Swiss firm in the usual way based on the proximity to cited firms' patent stocks worldwide.

The use of survey data combined with patent data has the advantage that we can measure a firm's innovation success with a variable measuring sales with innovative products, which is a better proxy for the commercial success of innovation activities than frequently used binary proxies or patent counts. In addition, we are able to control for important industry and firm-specific factors.

The new spillover measure is tested in the framework of an innovation equation in which we control, among other things, for absorptive capacity, appropriability, and competition conditions. The spillover proxy based on cited firms worldwide shows a positive and highly significant effect on innovation success. A statistically significant, positive effect is also found for a spillover variable that is based on cited firms from Switzerland only. In contrast, only a relatively weak association with innovation success could be found for a spillover measure that is calculated for all Swiss applicants (irrespective of whether these firms are cited). The marginal effect of the new measure is not only larger, it also measures the relationship more precisely than the traditional measure. In addition, the spillover effect is stronger for sales stemming from modified products as compared to new products.

The results for regional spillovers show that cited firms' knowledge stocks both in Switzerland and in European countries matter for the commercial innovation success. For spillovers stemming from the USA or Japanese firms, we do not find an effect, which might be due to localization of spillovers as well as to differences between the countries with respect to their technological orientation.

A further contribution of this study is that we analyze interactions between knowledge spillovers and the degree of competition in the product market. Although the competition-innovation relationship has been investigated extensively, we are lacking studies looking at the impact of competition on knowledge spillovers empirically at firm level. We found that an increasing number of principal competitors in the main sales market worldwide of the focal firm reduces the spillover effect from cited firms. This result indicates that spillover effects on innovation performance are at largest for firms that operate in niche markets or in oligopolistic structures. However, this effect can be traced back solely to innovators with new products as compared to only modified products.

The paper is structured as follows: In Section 2.2, the conceptual background, our 'new' spillover measure and the research hypotheses are presented. In Section 2.3, the specification of the empirical model and econometric issues are described. Section 2.4 describes the data that is used and Section 2.5 presents the results. Section 2.6 summarizes and concludes.

2.2 Conceptual Background

2.2.1 Knowledge Spillovers: Concept and Measurement

Overview

A crucial aspect of innovative activity is the generation of knowledge, which to some extent has the character of a public good. This gives rise to externalities ('spillovers') that are a central theme in the literature on innovation in industrial economics (see, e.g., Aghion and Jaravel 2015; Cohen and Levinthal 1989; Geroski 1995; Griliches 1979; Spence 1984).

A general though rather simplistic way to address this externality problem is to assume the diffusion of new private knowledge leading to a 'spillover pool of knowledge' from which other economic actors can draw information useful for their own innovative activities. A general formulation for the spillovers as a (weighted) sum of the knowledge capital of a firm's relevant technological environment that gives rise to a knowledge pool is given by the following expression (see Griliches 1979, 1992):

$$SO_i = \sum_j w_{ij} K_j ; i \neq j \tag{2.2.1}$$

for focal firm i, where K_j is the patent-based knowledge capital of firm j belonging to the relevant economic environment of the focal firm i; w_{ij} is a weighting variable to be further specified. On what should such a weighting variable be based? Broadly speaking, two distinct concepts of knowledge spillovers have been applied in literature (see De La Potterie 1997, for a review). According to the first one, spillover knowledge is related to flows of intermediate and/or capital goods and is assumed to be proportional to the value of the stream of goods between firms/industries (see, e.g., Wolff and Nadiri 1993). In the second concept, the weights in equation (2.2.1) are a measure of scientific and technological 'distance' among firms and industries (technological proximity; see, e.g., Bloom et al. 2013; Jaffe 1986) or of geographical distance (geographical proximity; see, e.g., Bloch 2013; Gust-Bardon et al. 2012). Here, we focus on measures of technological proximity.

The well-known Jaffe technological proximity measure between all firm pairings in a certain sample of enterprises takes the following form:

$$TECH_{ij} = \frac{T_i T_j'}{(T_i T_i')^{1/2}(T_j T_j')^{1/2}}; i \neq j \tag{2.2.2}$$

where T_i and T_j are vectors containing the shares of patents of each firm in each technological field; $T_i = (T_{i1}, T_{i2}, , T_{iF})$ for F distinct technological fields. The pool of technology spillovers of the focal firm i in year t is proxied by what we call 'spillover measure':

$$SPILL\_JAFFE_i = \sum_j TECH_{ij} K_j; i \neq j \tag{2.2.3}$$

where K_j is the knowledge stock of firm j.
A major limitation of studies using this traditional measure is that they only focus on sample firms, i.e., firm i and firm j must be necessarily in the same sample. Because the firm datasets very often only comprise firms from one country (and in the most famous studies only firms from the US), it is not possible to account for spillovers that might come from firms outside the focal country. Although spillovers have been found to be localized (see, e.g., Jaffe et al. 1993), in a globalized world it is most likely that there are still spillovers from foreign countries that are not negligible.

A New Spillover Measure: Technological Relevance and Foreign Spillovers

In this paper, we both restrict and at the same time expand substantially the pool of firms from which a focal firm in our sample can receive spillovers. As a result, we obtain a new measure that might have advantages compared to the traditional Jaffe measure as it takes into account technological relevance and foreign spillovers. The last point is especially interesting in the case of Switzerland for which we have firm-level data. The position of Switzerland in the innovation global landscape is quite strong and firms are acting globally. As a consequence, they are also searching for knowledge globally. Especially for a small country, in-sample spillovers might neglect a substantial part of incoming knowledge from foreign countries and/or from firms that are not in the sample. 'Technologically relevant' firms worldwide are defined as those firms whose patents are cited in the focal Swiss firm's patents (backward citations). We identified all firms that are cited by Swiss firms in their patent applications to construct the sample of firms that build the technologically relevant environment of a focal firm. We consider backward citations to be a good proxy for the technological relevance of patents for the citing focal firm because it is likely that a firm cites patents (or examiners assign citations to its patents) from firms that are active in similar industries, technological areas, etc.

Once we have identified the cited firms for each Swiss firm, we calculated the Jaffe proximity measures for $i = 1, ..., n$ sub-samples, where n is the number of Swiss patenting firms in our sample. Each sub-sample contains $1 + n_{cited\_i}$ firms where n_{cited_i} is the number of firms cited by Swiss firm i. Each of these sub-samples defines the technologically relevant environment for the respective focal firm. For the calculation of the spillover variable we use only the proximity measures between the focal firm and the n_{cited_i} firms in sub-sample i. As compared with the Jaffe measure the difference is that only those firms are taken into consideration for constructing the spillover variable whose patents (more precisely: at least 1 patent) have been cited in the patents of the focal firm (backward citations).[1]

2.2.2 Knowledge Spillovers and Innovation Performance

The relationship between knowledge spillovers and innovation performance is investigated in most extant studies in the framework of a patent equation which approximates a knowledge production function (see, e.g., Pakes and Griliches 1984) containing primarily R&D inputs and measures of knowledge spillovers based on patent or R&D stocks. The main idea is that knowledge spillovers may offer additional know-how to firms that are able to absorb such knowledge and combine it with in-house generated knowledge. Cohen and Levinthal (1989, 1990) demonstrated that knowledge spillovers can induce complementarities in R&D efforts and introduced the notion of absorptive capacity as the precondition for a firm to be able to exploit such spillovers. Hence, given a certain degree of absorptive capacity, the impact on innovation performance is expected to be positive in general, eventually mitigated by appropriability and/or competition factors (see below).

A positive effect of the R&D-based spillover variable on the number of patents has been already found in the seminal study of Jaffe (1986) for two cross-sections in 1973 and 1979 comprising 432 American firms. Peri (2005) also reported a positive impact of a patent-based spillover variable on the number of patents of 147 US regions in the period 1975-1996. In a recent study, Bloom

[1] Actually, we calculated the Jaffe proximity measure for all firm pairings in each subsample, i.e., the focal firm i and the n_{cited_i} cited firms in subsample i, and we eliminated the interactions between the cited firms n_{cited_i} themselves as we did not need them for the construction of the spillover variable for the focal firms.

et al. (2013) investigated the relationship between two patent-based technological spillover variables and innovation output measured by the number of patents and found positive effects of spillovers on patents for a panel of US firms for the period 1981-2001.

Furthermore, two European studies, one based on data for Italian firms and the second on data for German firms, investigated the impact of R&D-based knowledge spillovers on measures of innovation output other than patents. Cardamone (2010) examined the impact of technological spillovers for a panel of 1,203 Italian firms over the period 1998-2003. The results showed that the probability of introducing a product or process innovation is negatively correlated with technological spillovers, contrary to the findings of most other studies. Jirjahn and Kraft (2011) examined the effects of spillovers as measured by a binary variable for 'firm taking innovation ideas from observing competitors' on innovation output based on pooled data for 1022 manufacturing firms in Lower Saxony covering the period 1995 and 1997. They found that spillovers have a positive impact on the probability of introducing 'incremental' innovations but no effect on the probability of 'drastic' innovations.

Based on the above discussion of extant literature, we formulate the following hypothesis:

Hypothesis 1: There is a positive relationship between knowledge spillovers and innovation performance.

2.2.3 Localization of Knowledge Spillovers

The main idea is that geographical (spatial) proximity enhances the ability of firms to recognize and absorb external knowledge that is relevant for this firm's innovation activities by reducing the inherent uncertainty of identification of relevant knowledge (see, e.g., Audretsch and Feldman 1996). Of course, in a world in which geographically dispersed activities can be linked electronically, the importance of geographic location as a factor of knowledge creation may seem irrelevant. Nevertheless, many empirical studies confirm that geographical distance still plays a significant role for the degree of knowledge diffusion. In particular, this is the case for the transfer of tacit knowledge components (see, e.g., Gertler 2003). Empirical evidence on spatial proximity is often based on patent citations by comparing the geographical location of patent citations with that of the cited patents. Feldman and Kogler (2010) surveyed the relevant literature and they found that most empirical studies confirm that knowledge spillovers are localized.

However, only geographical proximity may not be sufficient for the existence of knowledge spillovers. As Feldman and Kogler (2010) emphasized, cognitive distance, proxied, for example, by the Jaffe technological proximity measure, is a further important factor which could enhance knowledge diffusion if the technological profiles are close enough to enable absorption and implementation of external knowledge. However, if the technological profiles are too similar, the generated spillovers may be of minimal added value and consequently would not positively contribute to the innovation performance of the focal firm. As already stated in the seminal paper of Jaffe et al. (1993), the disentanglement of the two effects is not easy if the focus is on spatial proximity because "there are other sources of agglomeration effects that could explain the geographic concentration of technologically related activities without resort to localization of knowledge spillovers" (p. 579).

The main result of Jaffe et al. (1993) based on citations of patents that were granted by the US patent office was that citations to domestic patents are more likely to be domestic and even more likely to come from the same state as the cited patents. Localization fades over time but slowly. In contrast, Li (2014) found that distance effects increase over time for the same age of citations; otherwise, this study also supports the localization hypothesis.

In a further paper, Jaffe and Trajtenberg (1999) found based on citations of patents granted by the US patent office to inventors in the US, the UK, France, Germany, and Japan with respect to spatial distance that patents whose inventors reside in the same country are 30% to 80% more likely to cite each other than inventors from other countries. Hence, the spillover localization tendency seems not only to occur in the US. The existence of localized spillovers has been challenged by Thompson and Fox-Kean (2005) substantially from a methodological point of view.[2] In a recent study, Murata et al. (2014) found based on a new distance-based test solid evidence supporting localization.

Further studies that support the localization hypothesis can be found in Peri (2005), based on patent citations for 113 European and North American regions over 22 years, Maurseth and Verspagen (2002), based on patent citations for European regions, and Fischer et al. (2009), based on high-tech patent citations in Europe.

[2]This has been the subject of the debate in the *American Economic Review* between Henderson et al. (2005) and Thompson and Fox-Kean (2005).

For our study we formulate the following hypothesis:

Hypothesis 2: Knowledge spillovers are stronger the smaller the geographic distance among interacting firms is, other things being equal.

In the case of Switzerland, we thus expect that spillovers from firms in Switzerland will show a stronger association with innovation performance than those from firms from other countries and spillovers from firms in Europe a stronger association than those from firms from other more distant regions.

2.2.4 Knowledge Spillovers and Competition

Contrary to the extensive theoretical and empirical literature on the relationship between competition and innovation performance (see, e.g., the seminal paper of Aghion et al. 2005), research is silent about a possible moderating effect of competition on the innovation effect of spillovers. Under the assumption that the amount of spillovers is directly and positively related to the innovation performance of a firm, one could formulate the following hypothesis about the moderating performance effect of competition: if a competitive situation generates a large amount of spillovers then the expected performance effect is presumably high (positive) and if a competitive situation generates few spillovers the expected performance effect is low (negative). However, even in this respect the literature is not definite. In a survey of theoretical literature, De Bondt (1997) refers indirectly to this non-linearity concluding as follows: "In strategic investment [...] more spillovers typically lower effort, unless other factors such as a not too competitive oligopoly (high degree of product differentiation, small number of rivals) render the leakage effect small and then the opposite tendency may apply" (p. 13).[3]

There are some investigations about the amount of spillovers generated in specific competitive situations. Zirulia and Lacetera (2010) develop a model in which high knowledge spillovers lead firms to soften incentives [of scientists for R&D] in order not to benefit competitors, but only when product market competition is high; in contrast, high spillovers positively affect incentives when competition is low, yielding a non-linear relationship between the degree of spillovers and

[3] It has to be remarked that in this approach the appropriability aspect is not separated from the knowledge aspect.

competition intensity.

With an agent-based simulation model, Wersching (2010) comes to the opposite results. He discusses the two views of Schumpeterian competition and their implications for innovation performance taking also knowledge spillovers into account. The simulation results show that a technological regime with many competitors in the product market is compatible with strong spillovers and in the case of only few competitors with weak spillovers. Given that the theoretical discussion remains inconclusive, the issue of the influence of competition on the innovation effect of knowledge spillovers has to be settled empirically. Thus, we are agnostic and formulate the following three-part hypothesis:

Hypothesis 3a: Competition enhances the effect of knowledge spillovers on innovation performance.

Hypothesis 3b: Competition reduces the effect of knowledge spillovers on innovation performance.

Hypothesis 3c: Competition does not affect the effect of knowledge spillovers on innovation performance.

2.3 Model Specification and Econometric Issues

2.3.1 Model Specification

The usual framework to study the impact of technological knowledge and knowledge spillovers on innovation performance at the firm level is the knowledge production function which models the relationship between innovation input and innovation output (see for a standard model Crépon et al. (1998) and Cohen (2010) for a survey of this literature). We formulate this relationship as a function between the sales of innovative products (LINNS) (that includes sales with new and significantly modified products), i.e. a measure of innovation success, and the knowledge capital (LK) as well as knowledge spillovers (LSPILL) that contribute to this success (see Ramani et al. 2008, for a similar approach):

$$LINNS_{it} = \alpha_0 + \alpha_1 LK_{it-1} + \alpha_2 LSPILL_{it-1} + \alpha_3 X_{it-1} + e_{it} \tag{2.3.1}$$

where

$X_i = \{D_i; IPC_i; INPC_i; NCOMP_i; APPR_i; LEMPL_i; HQUAL_i; FOREIGN_i;$ industry dummies; year dummies$\}$;

for firm i, year t. Thus, the total impact of knowledge on firm output is measured by $(\alpha_1 + \alpha_2)$, the sum of the effects of a firm's own knowledge capital and the knowledge obtained by spillovers from enterprises of a firm's technologically relevant economic environment. We control for demand conditions (D), competition conditions (IPC; INPC; NCOMP), appropriability (APPR), the degree of absorptive capacity that is proxied with the share of highly qualified employees (HQUAL), firm size (LEMPL), foreign ownership (FOREIGN), industry affiliation and reference year (see Table 4.A.1 for the exact definition of the variables). Controlling for appropriability and absorptive capacity is particularly relevant in our approach of firms perceiving spillovers that are based on patent citations as measures of technological linkages among firms. Competition conditions are measured by a structural variable (number of main competitors in the relevant product market worldwide).

2.3.2 Econometric Issues

We estimate the reduced form in (2.3.1) with Generalized Least Squares (GLS). Standard errors are heteroscedasticity and cluster robust. Reverse causality is not a concern in this setting since all covariates are lagged by one period.[4] Although we control for absorptive capacity and the existing knowledge stock, we are not able to include all firm-specific factors that are relevant to enable a firm to absorb spillovers from the technological environment, e.g., we do not observe management quality. In additional estimations that are detailed in 2.5.4, the potential endogeneity of the spillover variable (LSPILL) is addressed by using additional lagged levels and differences of the focal variable as instruments. We have to note that in some of the empirical spillover literature, own knowledge capital rather than the spillover variable is assumed to be endogenous (e.g., in Lychagin et al. 2016). We mainly follow Bloom et al. (2013) and focus on

[4] In fact, the covariates are lagged by three years. This is due to the survey data we use which is only available for each third year, see next section.

the spillover variable, but in Table 2.B.4 we also present a specification where we treat both variables as endogenous.

2.4 Data

2.4.1 Swiss Innovation Panel

The data stems from 6 waves of the Swiss Innovation Survey conducted by the KOF in the years 1996, 1999, 2002, 2005, and 2008.[5] The surveys are based on a disproportionately stratified random sample of firms with more than 5 employees (in full time equivalents) covering the industries of the manufacturing, construction and (commercial) service sector. The sample stratification refers to 2-digit industries and within each industry to three industry-specific firm size classes. The investigation at hand only uses data for manufacturing firms with patent applications with 264, 316, 328, 332, and 304 observations for the years 1996, 1999, 2002, 2005 and 2008 respectively. The resulting panel dataset is highly unbalanced. Due to missing values for model variables we end up with 640 observations in the pooled version (see Table 2.A.3 for the composition by industry, firm size class and year of the sample used in the econometric estimations; Table 2.A.2 for descriptive statistics; and Table 2.B.1 for the correlations between the model variables).

2.4.2 Patent Data

Annual information about patent applications comes from PATSTAT (EPO 2013) and the Derwent World Patent Index (WPI) by Thomson Reuters.[6] Based on the number of patent applications, we calculated patent stocks as proxies for knowledge stocks for each firm and year using the perpetual inventory method and a depreciation rate of 15% (see Hall et al. 2010):

[5] The questionnaire of the survey, which resembles closely the "Community Innovation Survey", is available at www.kof.ethz.ch in German, Italian, and French. In the estimations, we use three-year lags for all variables except for the dependent variable that comes from the 2011 survey.

[6] We conducted several rounds of names matching: First, we used all patent applicants from Swiss applicants from WPI between 1990 and 2010, cleaned the applicants' names and firm names, and matched the cleaned applicants' names with firm names from the Innovation Survey automatically with a matching software. Afterwards, we checked the results manually. We also searched each firm name from the panel in ESPACENET and PATSTAT to get as many as possible patent applications. At the end, all matched patent applications we found were compiled in one dataset and checked once again. For the analysis here, we use patent families rather than single applications. Families comprise multiple applications of the same invention in different countries. Thus, they better reflect inventions than single patent applications (Martínez 2011; OECD 2009).

$$K_{it} = (1 - d)K_{it-1} + R_{it} \qquad (2.4.1)$$

where K_{it} is the patent capital of firm i in t, d the depreciation rate, and R_{it} new patent applications in t. The initial value is calculated as follows:

$$K_{i0} = R_{i0}/(d + g) \qquad (2.4.2)$$

The growth rate g is calculated from the 10-year average growth rate at 2-digit industry level for patent applications before 1990.[7] Table 2.A.4 shows the calculated average patent capital by industry, firm size class and year. Chemicals, machinery and electronics/instruments are the industries with the largest patent stocks reflecting their high level of patenting activities.

The patent data also entails information about the technological fields (IPC code) at different levels of aggregation. We use the subclass level with four digits (for further explanations, see WIPO 2014) yielding 617 subclasses for the calculation of the Jaffe measure of technological proximity (see equations (2.2.2) and (2.2.3) in Section 2.2).

2.4.3 The EEE-PPAT Tables with Names of Applicants

We identified all cited firms with the EEE-PPAT table that contains cleaned and harmonized names of applicants.[8] First, we searched for patent applications that are cited by patents assigned to a Swiss applicant. These patent applications were matched with the 'person' table from PATSTAT and then matched with names and IDs from the EEE-PPAT table. In sum, we found 125,449 distinct firms that are cited by Swiss firms from our sample (including self-citations). The distribution of the number of cited firms is quite skew. In fact, 10% of the firms account for about 75% of all backward links. 50% of the firms have less than 31 backward links, whereas 1% of the firms have more than 2,460 links.

In the next step, we collected all patent applications for each cited firm in PATSTAT. This enabled us to calculate the patent stocks of cited firms in the same way we did it for the Swiss

[7] The reason for using industry-level information is that we did not match older patent applications before 1990. The sector assignment of patent applications necessitated the use of concordance tables, in our case that by Lybbert and Zolas (2014).

[8] See Du Plessis et al. (2009); Magerman et al. (2009); Peeters et al. (2009) for a description of the harmonization routines.

firms using the perpetual inventory method.[9] We also assigned technological fields at subclass level to each patent application starting from the year 1995. We ended up with N datasets for the N sub-samples described above. For each sub-sample, we calculated the firms' share of patents in the underlying subclasses (pooled over all years). Each dataset has $F \times (1 + n_{cited\_i})$ observations. Finally, we calculated the spillover measures using a programming loop over all datasets. The final measures for the Swiss firms were then assigned to the firm IDs in the innovation survey.

2.4.4 Spillover Variables for Different Regions

Based on formula (2.2.2), we first calculated the spillover measure that takes into account all backward citation links (see Table 2.A.5 for the average number of backward citations of Swiss firms by industry and by firm size class; Chemicals, Machinery, Electrical Machinery and Electronics/Instruments, which are the most innovative industries, show the highest number of citations). In a further step, we looked at different geographical areas separately, i.e., we calculated the measure only based on cited firms that belong to certain regions as identified by the person country codes of the patent applicants. As main regions of interest, we chose Switzerland (as home-base), 'Europe' (i.e., all European countries except for Switzerland), the United States and Japan. The United States and Japan are chosen because of their economic and technological importance and because of their importance as patentees that makes them a potential technological source. For each region r, we get $i = 1, ..., n$ sub-samples with $1+n_{cited\_i\_r}$ firms where $n_{cited\_i\_r}$ is the number of firms in region r cited by Swiss firm i (see Table 2.A.4 for the calculated average patent capital of the cited firms by regions).

For comparison with the measure based on cited Swiss firms, we also calculated the spillover measure in the usual way where we only take into account Swiss applicants irrespective of whether they are cited or not (formula (2.2.3)).[10]

[9] As we can directly query the EEE-PPAT IDs in PATSTAT, we were able to retrieve patent applications up to 1971.

[10] However, in contrast to the 'traditional approach', we take into account all Swiss applicants and not solely Swiss applicants that are part of the sample.

2.4.5 Self-citations

From formulas (2.2.1) to (2.2.3), it is immediately clear that backward links that are based on self-citations must be excluded. Otherwise, our measure would not measure incoming external knowledge spillovers properly. More severely, the knowledge capital of a focal firm would enter the right-hand side of the regressions twice: First, as a focal firm's knowledge stock and, second, as weighted external knowledge stock through the spillover measure.

Excluding self-citations is involved because we have to deal with datasets with different firm identifiers: The survey data uses other identifiers than the EEE-PPAT table. Therefore, we cannot simply match the two data sources based on firm IDs. However, we can identify 'matching' firms in the respective datasets based on the patent applications they have in common. Concretely, we used all backward citations we could find for the Swiss firms (citing firms) and matched both the cited and citing patent applications with IDs from the EEE-PPAT table. Afterwards, we deleted all backward links where the cited patent applicant and the citing patent applicant have the same firm ID from the EEE-PPAT table in common in order to eliminate systematically all links between entities that might belong to the same company or are in any kind of judicial relationship.[11] The number of backward links then drops to 122,629.

2.4.6 Potential Biases and Problems with Patent Data

In the European patent system, most of the citations are added by patent examiners rather than by applicants or inventors (see Criscuolo and Verspagen 2008). Nevertheless, many authors use citation counts as – perhaps noisy – proxies for knowledge flows. Schoenmakers and Duysters (2010) argue that inventors might not bother to include a citation and that they might simply forget to include a citation, or even deliberately not include a citation for strategic reasons. Overall, they conclude that particularly with respect to the European Patent Office also non-inventor citations might indicate knowledge flows very well. Duguet and MacGarvie (2005) analyzed to what extent survey responses to the French innovation survey on R&D outsourcing, external R&D, cooperative R&D and other technology sources can predict backward and forward citations. They found support for the claim that patent citations are associated with technology flows as identified from the survey questions for some, but not all, channels. In contrast to citations that refer to economies that are more integrated with the French econ-

[11] This might apply to some foreign subsidiaries.

omy, citations to US inventors are associated with technology acquisition through more indirect means such as equipment purchases. Roach and Cohen (2013) did a similar exercise for knowledge flows from US research institutes to firms and found that citations reflect knowledge flows through channels of 'open science', but not through contract-based relationships.

In this paper, we assume that examiners add citations that reflect their expert opinion covering existing patented knowledge on the topic in question. We do not see any reason why applicants should not also have perceived the same knowledge as examiners, even if they have not reported it in their applications. Consequently, we assume that citations (including examiners' additional citations) can be at least used to identify firms that are relevant for a focal firm from a technological point of view.[12] In additional estimates, we investigated the influence of examiner citations on the robustness of our results. Using only citations that were added by applicants does not considerably change the elasticity of the spillover variable for all regions (0.099 versus 0.093, see Table 2.B.2 and the discussion in Section 2.5.4). Thus, our estimates are quite robust with respect to the distinction between citations that were added by the examiner or the applicant or solely by the applicant.

Our results might be confronted with some other potential biases that arise from different aspects of the underlying data and the patent system. The latter are discussed in De Rassenfosse et al. (2013) and Bacchiocchi and Montobbio (2010). First, results might be subject to an institutional bias when patents are used that are from countries with different patent systems. However, this problem can be mitigated by using patent families as we did.[13] Second, there might be a geographic bias as applicants tend to file in their home patent offices and examiners tend to cite patents from their home offices. However, we reduce the possibility of this bias by avoiding looking at single patent offices. De Rassenfosse et al. (2013) found that small countries such as Belgium, the Netherlands, and Switzerland first file their patents at the European Patent Office. Thus, this kind of bias can be avoided in case of European cited firms. Problems might arise, if, for example, a US firm only applies in the US but not in Europe and the respective patent is not cited by a Swiss firm only because it is not applied for in Europe. We assume that 'technologically relevant' patents are mostly filed also at the European Patent Office (as the most important patent office beneath the USPTO and the JPO) even if the applicant is from the US.[14] Moreover, patent families that comprise a large number of patents that have been

[12] We do not attempt to capture knowledge flows with backward citations in this paper.

[13] In fact, we use families for both 'cited' and 'citing' patents.

[14] A large number of the cited firms are US firms, namely 39,437 compared 32,778 European firms. We also want to emphasize that a home bias with respect to USPTO citations (the citation practices are different from

applied for internationally are more valuable (Harhoff et al. 2003). Therefore, relevant patent families should comprise patent applications in multiple geographical jurisdictions. A final argument against a geographical bias is that we only look at backward citation links and not at the number of backward citations. Once a foreign firm has received one backward citation, it is taken into account in our analysis.[15]

A further bias might arise from including backward citations to patents that were applied for or granted a long time ago. However, we argue that we are not interested in the cited invention per se, but rather in the general technological relevance of the cited firm. If a patent cites an invention that was made a long time ago, the cited invention or at least the firm behind it should still possess technological characteristics that could make it a potential spillover source otherwise it would not have been cited by the focus firm.

There might be also concerns that our results are driven by firm size and the Chemical and Pharmaceutical industry (the largest firms with the largest number of patent applications can be found here). However, inclusion of these firms is essential as they might be important spillover sources for smaller firms in Switzerland and their knowledge capital might affect the innovation performance of other firms through our spillover measure.[16]

2.5 Results

2.5.1 Basic Model and Comparison of Spillover Measures

Columns 2 and 5 in Table 2.A.6 show the estimates for the basic model for LINNS based on the spillover variables LSPILL and LSPILL_CH according to equation (2.2.3). LSPILL is based on all backward links, whereas LSPILL_CH only refers to cited Swiss firms. Both the elasticity of the knowledge capital and the spillover variables are positive and statistically significant (columns 2 and 5). For SPILL_CH, an increase by 1% of a firm's knowledge capital is associated with an increase by 0.123% of the sales of innovative products. The respective elasticity for the spillover variable is 0.099 (0.094 for LSPILL). Thus, the joint effect of own and spillover patent capital amounts to 0.222 (0.223 for LSPILL), i.e. a change of 1% of the joint knowledge capital

the European patent system) does not matter as we only look at citations Swiss firms made rather than citations US firms made.

[15] In Table 2.B.2, columns 5 and 6, we show estimates when the spillover measure is additionally weighted with the share of backward citations.

[16] In the regressions, we control for firm size that is strongly correlated with the number of backward citations and the number of patents.

is related to a change of 0.222% of innovative sales.[17] The positive sign of the spillover variable confirms hypothesis 1.

We compare the estimates for the new citation-based measure referring to cited Swiss firms (LSPILL_CH) with the estimates for a standard Jaffe spillover variable based on patent stocks of all Swiss firms with patents (LSPILL_ALL; Table 2.A.6, column 1), irrespective of whether they are cited or not. The coefficient of the spillover variable is 0.053, i.e., much smaller than that for the new measure, and statistically significant at the 10% test level. We interpret this result as evidence that the new spillover variable identifies more relevant external knowledge as shown by a substantial larger contribution (larger elasticity) of spillovers to a firm's innovation success. Hence, the better performance of the new measure is presumably due to the identification of firms that are really technologically relevant for the focal firm as identified with backward citation links reflecting links to external knowledge that the focal firm anticipated when generating its inventions.

Further variables that show - as expected - positive and statistically significant coefficients at the usual test levels are the measure for absorptive capacity (HQUAL) and the measure for firm size (LEMPL). The coefficient of the appropriability variable (APPR), a further relevant control variable, is positive but not significant for LINNS.

In columns 3 and 4 and 6 and 7, respectively, we investigate the new spillover variables for the sales with 'new' (LINNS_N) and 'significantly modified' (LINNS_M) products separately. This distinction captures more radical vs. more incremental innovations. The results show that spillover-related patent capital is significantly more important for modified products than for new products. The elasticity is 0.113 for LINNS_M as compared to 0.056 for LINNS_N for spillovers from cited Swiss firms and 0.098 versus 0.073 for spillovers from all cited firms worldwide. Moreover, own patent capital is insignificant for modified products, but highly significant for new products. This indicates that incremental innovation success is more dependent on external knowledge ('open innovation'), whereas radical innovation success is more related to exploitation of own knowledge resources. Indeed, APPR shows a positive and significant coefficient for the sales with new products, thus supporting this presumption.[18]

[17] In a recent study based on data for several OECD countries for the period 1974-2002, Acharya (2015) estimated an average elasticity of intra-industry R&D spillovers (with respect to labor productivity) of 0.071, which is of the same magnitude as our estimates at firm level.

[18] Our results are in line with Jirjahn and Kraft (2011) who have found that spillovers do not stimulate drastic innovations, although they solely rely on survey data and the dependent variable and spillover variable are therefore specified differently.

2.5.2 Basic Model and Competition Effects

Table 2.A.7, columns 1 to 3 shows the estimates of the basic model expanded by the interaction term between the overall spillover variable LSPILL and the competition variable NCOMP that measures the number of principal competitors on the main product market. The coefficient of NCOMP becomes significantly positive in the estimates for LINNS_N and remains insignificant for LINNS_M. The coefficient of the interaction term is negative and statistically significant at the usual test level. This means that the effect of spillovers on the commercial success of innovations is significantly lower in markets with a larger number of competitors. This negative effect can be traced back primarily to sales with new products (see column 2) and is a hint in favor of Hypothesis 3b. Obviously, more competition on the product market increases the need to innovate more radically, but reduces the contribution of spillovers to innovation success with new products. In face of stronger competition, radical innovators might also be careful that own knowledge does not leak out to rivals; this explains the positive sign of the appropriability variable in the estimates for LINNS_N.

In columns 4 to 6, we specify the competition effect in an alternative way by interacting the spillover variable with a dummy variable that takes on value 1 if the number of competitors is larger than 15, this being the cut-off value from where on competition matters (see the significantly positive coefficient of the dummy variable NCOMP>15 in columns 4 and 6). The negative and statistically significant coefficient of the interaction term confirms the previous results and shows that increasing the number of competitors above 15 decreases the spillover effect on innovation performance. Again, this result can be primarily traced back to innovation success with new products. These results support Hypothesis 3b with respect to the number of competitors as a measure of market concentration.

There are two possible interpretations of the finding that the effect of spillovers is weakened when firms are operating in markets with many competitors (polypolistic markets). One possible explanation for this result refers to the size of the knowledge capital stock of the cited firms.[19] In polypolistic markets, firms lack the financial means for comprehensive investments in R&D and consequently, knowledge advancements are weaker and the average knowledge capital stock is likely to be lower than in markets with less competitors. Hence, fewer spillovers

[19] However, the spillover measure also includes the technological proximity measure, which is multiplied with the size of the knowledge capital stock. The proposed explanation only refers to the knowledge capital stock. More in-depth analysis would be necessary to include the proximity measure into the explanation of the observed facts.

are generated and their effect on innovation performance is lower. In markets with few R&D active competitors, the firms' knowledge stocks are likely to be higher on average, hence, more spillovers are generated and their effect on innovation performance is expected to be larger. The second interpretation refers to a kind of 'business-stealing effect' as described in Bloom et al. (2013) and is more related to the total knowledge stock in a market – that is increased by a larger number of R&D active firms – rather than the average knowledge stock per firm: Firms benefit from competition on the technological market if the social benefits arising from R&D spillovers exceed the costs of knowledge leakage, but they suffer from competition on the product market at the same time so that the compound effect of competition and spillovers on performance measures is negative.

On the whole, it appears that knowledge spillovers contribute disproportionately stronger to innovation success in more concentrated markets for a given level of appropriability and absorptive capacity. This is particularly the case when innovating firms pursue strategies of high degree of innovativeness.

2.5.3 Regional Effects

As already described, we calculated separate spillover variables based on backward citation links to Swiss firms only (LSPILL_CH), to European firms (LSPILL_EU), US firms (LSPILL_US) and Japanese firms (LSPILL_JP), respectively. We inserted all four regional spillover variables in the LINNS equation and estimated the model once again (Table 2.A.8, column 1). In a further step, we inserted the four regional spillover variables separately in the innovation equation and estimated four different models (columns 2 to 5). The estimates with all four spillover variables show that only the coefficient of the spillover variable from other Swiss firms is positive and statistically significant. Thus, the overall spillover effect can be traced back mainly to spillovers from other Swiss firms, the geographically nearest economic environment of a Swiss enterprise. The separate estimates for each regional variable confirm this finding and yield the additional insight that European firms also contribute to knowledge spillovers of Swiss firms, but to a smaller extent than Swiss firms (0.063; column 3). The coefficients of the spillovers from US and Japanese firms are negative and statistically insignificant.[20] These results support Hypothesis 2 and they are in accordance with the findings of recent studies for the US (Li 2014, ; based on

[20] A distinction between new and modified products did not yield any further insights. Therefore, results are not shown here.

citations for the period 1980-1997) and six large industrial countries for the period 1980-2000 (Malerba et al. 2013). Although American and Japanese firms possess of quite large patent stocks on average, spillovers from these stocks result in smaller effects than the much smaller stocks of Swiss and European firms. A possible explanation is a large technological distance between Swiss firms and firms from USA and Japan. Hence, the regional effect might be strengthened by the technological proximity effect.

2.5.4 Robustness Tests and Further Estimations

Robustness of the Spillover Effect

We conducted three robustness tests with respect to the effects of the spillover variable and the competition effects. The tests refer to (a) the exclusion of examiner citations, (b) the exclusion of non-profit organizations, and (c) the consideration of weights for the backward citations. Table 2.B.2 (columns 1 and 2) shows results where we only include links based on citations made by applicants and exclude citations added by examiners. The results are robust. The EEE-PPAT table contains sector affiliations of the patent applicants. It has to be mentioned that not all patent applicants are private firms although they are by far the majority.[21] In columns 3 and 4, we consider only spillovers from profit-oriented firms (other institutions were excluded before calculating the proximity measures). Again, the results are robust.

In columns 5 and 6, the spillover measure is weighted with the share of backward citations (i.e., the number of backward citations that occur between a Swiss firm and a cited firm relative to the total number of backward citations a Swiss firm made). Obviously, the relative number of backward citations has an influence on the magnitude of the effect of spillovers (as measured in this study) on innovation performance. The elasticity of the spillover measure becomes larger, but remains in the same range of magnitude. Thus, our findings without weighting are rather conservative, the elasticities of the spillover measures displaying a kind of lower bound. With respect to the competition effect, the interaction effect with competition is supported in the case of (a) and (b) (see column 2 and 4) but not in (c) (column 6).

[21] We can detect 118,373 private firms, 5,551 non-profit organizations, and 1,598 universities that were cited by Swiss applicants. Individuals can also apply for patents, but they were excluded from the analysis from the beginning.

Robustness of the Econometric Specification

In further estimations, we check the robustness of our results with respect to different econometric models and the possible endogeneity of model variables. Wooldridge (2010, pp. 70) considers simple proxy variable solutions in order to eliminate omitted variable bias. He uses the lagged dependent variable in order to proxy for unobserved heterogeneity. However, this procedure is only valid in the cross-section (provided that one lag of the dependent variable is available). We therefore eliminate the time dimension from our data, introduce the lagged dependent variable $LINNS_{i,-1}$ (all independent variables were lagged before eliminating the time dimension) and estimate Ordinary Least Squares. We assume that the unobserved heterogeneity q_i can be approximated by the lagged dependent variable. This means that:

$$q_i = \beta_0 + \beta_1 LINNS_{i,-1} + u_i \tag{2.5.1}$$

We assume that u_i has zero mean and we are confident that u_i is not significantly correlated with the spillover variable $LSPILL_{i,-1}$ and the proxy for competition in equation (2.3.1), which are the main variables of interest.[22] If this holds then α_3 and the coefficient of the competition variable $NCOMP_{-1}$ in equation (2.3.1) are unbiased and measure the effect of spillovers and competition on the innovation performance of firms in the cross-section. Why do we think that u is uncorrelated [$Cov(LSPILL, u) = 0$ and $Cov(NCOMP, u) = 0$]? Given our comprehensive control vector including the lagged dependent variable it is hard to think of lacking important information that is strongly correlated with spillovers and/or the number of principal competitors. Of course the lagged dependent variable is endogenous, however, since we are not interested in the marginal effect of this variable, this is of no concern to the empirical estimation strategy described here. Although the spillover effect decreases by about 21% (see Table 2.B.3), it is still highly significant and the results are robust even when accounting for unobserved heterogeneity in this simple way.

In addition, Table 2.B.4 shows Blundell-Bond estimates that are panel-consistent dynamic GMM estimates also including the lagged dependent variable as regressor (Arenallo and Bover

[22] The relationship we are interested in can be formulated by $LINNS_i = \gamma_0 + \gamma_1 q_i + \gamma_2 LK_{i,-1} + \gamma_3 LSPILL_{i,-1} + \gamma_4 X_{i,-1} + \epsilon_i$.

1995; Blundell and Bond 1998).[23] The estimator uses further lags and lagged differences of $LINNS_{i,t-1}$ as instruments in a differenced equation and in a level equation, respectively. Additionally, we account for potential endogeneity of LSPILL by introducing the lagged difference and the lags $LSPILL_{i,t-2}$ and $LSPILL_{i,t-3}$ as instruments in the system estimation (column 2). In column 3, we also include instruments for LK in the same way.[24] At the bottom of the table, we report the Sargan test statistics and tests on zero autocorrelation. The Sargan test on valid overidentifying restrictions cannot be rejected for the specifications in column (2) and (3). The tests on zero autocorrelation show that the errors are serially uncorrelated as we cannot reject at order 2. Although prior innovation success does not show any statistically significant association with current innovation success once we account for endogeneity of prior innovation success, the spillover effect remains highly significant which supports our baseline results. Interestingly, in the additional estimations we provide in Table 2.B.4, the coefficient of patent capital is most often insignificant - quite in contrast to our baseline results. Lychagin et al. (2016) also get insignificant coefficients for own knowledge capital in some specifications when accounting for spillovers. Although not explicitly discussed in literature so far, this might point to simultaneous relationships between knowledge capital and knowledge spillovers and might indicate the need to estimate a simultaneous equation model, e.g., a GMM Three-Stage Least Squares model as described in Wooldridge (2010).[25]

2.6 Summary and Conclusions

In this paper, we contribute to literature in three ways: First, we examine the impact of knowledge spillovers as measured by a patent-based proximity measure on innovation success. Second, we propose a new measure that extends the traditional Jaffe spillover measure; it uses backward citation links to identify the firms to which a focal firm is technologically exposed. Third, we

[23]The Blundell-Bond estimates turn out to be more efficient in our case than the Arenallo-Bond estimates (Arellano and Bond 1991; Blundell and Bond 1998). The method is similar to the one applied in Lychagin et al. (2016) (see Table VI in their paper) although they treat only own knowledge capital and lagged productivity (their lagged dependent variable) as endogenous.

[24]We do not lag the independent variables that we include in the regressions from the outset as in the baseline models. The reason is that the estimator uses lags and lagged differences for the endogenous variables and that we have to reject the test on joint significance of regressors when using lagged regressors from the outset in this model.

[25]More concretely, spillovers might affect LINNS mainly through knowledge capital as external knowledge might be absorbed into own knowledge so that the own knowledge stock increases which affects the sales with innovations positively. Preliminary results, not shown here, seem to support the presumption that knowledge capital that affects innovation success is partly determined by spillovers, but a more detailed analysis is left for future research.

investigate the performance effects of spillovers in markets with different degrees of competition. Based on a comprehensive data set comprising firm-level survey information for a representative panel of Swiss firms and patent information for all firms worldwide with patents that have been cited by Swiss firms, we found that (a) the proposed new spillover measure shows a positive and significant effect of knowledge spillovers on innovation success as measured by the sales share of innovative products; (b) spillovers are more important for innovation success with modified products (incremental innovations) as compared to new products (radical innovations), while a firm's own patent capital is more important for success with new products than with modified products; (c) the knowledge spillovers are localized and concentrated primarily in Switzerland and to a smaller extent in Europe; and (d) market competition is important for the innovation effects of spillovers, but only with respect to radical innovation success.

With respect to competition, we found that firms in markets with many competitors do not benefit from spillovers, while firms in markets with few competitors (less than 15) benefit more from spillovers, but only with respect to firms that innovate with new products. This result indicates that spillovers are more important for Swiss firms that operate in niche markets (e.g., measuring instruments) or in typical R&D intensive, oligopolistic markets (e.g., pharmaceuticals). It reflects exactly the innovation strategy of many Swiss firms as it is investigated and discussed in previous studies (see, e.g., Arvanitis 1997; Arvanitis and Hollenstein 1996). However, with respect to the direct spillover effect, firms with a higher level of innovativeness draw on own accumulated knowledge to a larger extent than on external knowledge from spillovers and try to prevent knowledge leakage to rivals.

From a theoretical point of view, a possible mechanism for explaining our finding is as follows: intensive competition as indicated by the presence of many principal competitors might reduce the financial opportunities to invest in R&D. As internal R&D contributes to the absorptive capacity that is needed for the exploitation of external knowledge, the lack of R&D investments tends to reduce the performance effects of spillovers.

It is a limitation of this study that we only consider spillovers from patenting firms. If firms do not patent their inventions, they might choose other means of knowledge protection, such as secrecy, first-mover advantages, etc. It is likely that, for example, 'secrecy' leads to lower knowledge externalities, but the extent of spillovers from other strategic appropriability mechanisms is unknown. Cohen et al. (2002) suggest that R&D spillovers are significantly greater in industries and countries where appropriability is low, notwithstanding the relative effectiveness of particular mechanisms. Future investigations could shed light on the spillover effects of different appropriability mechanisms, but they are not subject of the present study. A further limitation is that it refers to one country only. The matching of firm survey data with patent data for several countries with different technological profiles would enable researchers to test the citation-based spillover measure on a wider basis and gain additional insights with respect to the role of knowledge spillovers in the innovation process.

Finally, a limitation may lie in the measurement of the numbers of competitors that is a rather crude measure from the survey that applies to the overall competitive environment. A refinement, e.g., by looking at the number of competitors in different technological classes could lead to more fine-grained results. With respect to the interaction effect that we find for spillovers and competition, a theoretical model would clearly help to understand the mechanisms from a conceptual point of view. This paper is a first attempt to understand the mechanisms between competition, spillovers, own existing knowledge and innovation success from an empirical point of view. Further progress in this area might depend on finding suitable instruments for spillovers and on further disentangling the relationships between the focal variables applying appropriate econometric methods.

Appendix

2.A Tables

Table 2.A.1: Description of Variables

Variables	Definition
LINNS	Sales of innovative (new + significantly modified) products; natural logarithm
LINNS_N	Sales of innovative products that are new; natural logarithm
LINNS_M	Sales of innovative products that are significantly modified; natural logarithm
D	Expected demand at the product market; five-level ordinal variable (1: very weak demand development; 5: very strong demand development)
NCOMP	Number of competitors at the main product market; five-level ordinal variable (1: up to 5 competitors; 2: 6 to 10; 3: 11-15; 4: 16-50; 5: > 50)
APPR	Easiness of copying innovations; five-level ordinal variable (-1: very weak copy easiness; -5: very strong easiness)
LEMPL	Number of employees in full time equivalents; natural logarithm
HQUAL	Share of employees with tertiary level education
FOREIGN	Foreign-owned; binary variable: 1: yes; 0: no
LK	Knowledge capital based on patents; natural logarithm
LSPILL_ALL	Knowledge spillover based on interaction with all Swiss applicants that have at least 1 patent; natural logarithm
LSPILL	Knowledge spillover based on interaction with all applicants whose patents have been cited by the focus firms (backward citations); natural logarithm
LSPILL*NCOMP	Interaction term of LSPILL with NCOMP
LSPILL_CH	LSPILL based on backward citations only of Swiss applicants; natural logarithm
LSPILL_EU	LSPILL based on backward citations only of European applicants; natural logarithm
LSPILL_US	LSPILL based on backward citations only of US applicants; natural logarithm
LSPILL_JP	LSPILL based on backward citations only of Japanese applicants; natural logarithm
LSPILL_APP	LSPILL based on backward citations filed by the applicant (excluding those added by examiners); natural logarithm
LSPILL_FIRMS	LSPILL based on backward citations only of applicants that are private corporations; natural logarithm
LSPILL_BACK	LSPILL based on backward citations, weighted with the share of backward links cited by a firm; natural logarithm

Table 2.A.2: Descriptives

Variable	Mean	Std. Dev.	Min	Max
$LINNS_t$	16.698	1.678	10.597	22.585
D_{t-1}	3.491	0.811	1	5
$NCOMP_{t-1}$	2.17	1.183	1	5
$LEMPL_{t-1}$	5.218	1.309	0.693	9.952
$HQUAL_{t-1}$	22.878	15.549	0	86
$FOREIGN_{t-1}$	0.236	0.425	0	1
$APPR_{t-1}$	-2.459	1.093	-5	-1
LK_{t-1}	2.152	1.063	0.462	7.429
$LSPILL$	3.773	2.187	0	8.327
$LSPILL\_CH_{t-1}$	2.157	2.141	0	7.844
$LSPILL\_EU_{t-1}$	1.895	2.475	0	8.104
$LSPILL\_US_{t-1}$	0.957	2.249	0	8.388
$LSPILL\_JP_{t-1}$	0.327	1.45	0	8.414
$LSPILL\_APP_{t-1}$	3.735	2.257	0	8.64
$LSPILL\_FIRMS_{t-1}$	3.697	2.192	0	8.236
$LSPILL\_BACK_{t-1}$	0.986	0.882	0	5.037

Table 2.A.3: Composition of the Dataset by Industry, Firm Size Class and Year (Number of Firms)

Industry	N	in %
Food, beverage	12	1.9
Textiles	12	1.9
Clothing. Leather	1	0.2
Wood processing	9	1.4
Paper	10	1.6
Printing	10	1.6
Chemicals	50	7.8
Plastics, rubber	27	4.2
Glass, stone, clay	17	2.7
Metal	12	1.9
Metal working	46	7.2
Machinery	194	30.3
Electrical machinery	66	10.3
Electronics, instruments	130	20.3
Vehicles	14	2.2
Watches	9	1.4
Other manufacturing	21	3.3
Firm size		
Small (5-49)	135	21.1
Medium-sized (50-249)	342	53.4
Large (> 250)	163	25.5
Year		
1999	74	11.6
2002	120	18.8
2005	136	21.2
2008	156	24.3
2011	154	24.1
Total	640	100

Table 2.A.4: Patent Capital per Firm by Industry and Firm Size Class

Industry	
Food, beverage	55.81
Textiles	7.83
Clothing. Leather	6.12
Wood processing	2.02
Paper	5.58
Printing	2.94
Chemicals	104.29
Plastics, rubber	6.72
Glass, stone, clay	6.21
Metal	8.75
Metal working	6.24
Machinery	20.85
Electrical machinery	66.14
Electronics, instruments	22.56
Vehicles	5.21
Watches	4.04
Other manufacturing	3.42
Firm size	
Small (5-49)	3.80
Medium-sized (50-249)	11.63
Large (> 250)	85.20
Total	28.66

Table 2.A.5: Number of Backward Citations per Firm of Swiss Firms by Industry and Firm Size Class

Industry	
Food, beverage	58.43
Textiles	5.57
Clothing. Leather	3.90
Wood processing	0.33
Paper	1.80
Printing	3.74
Chemicals	100.86
Plastics, rubber	6.96
Glass, stone, clay	5.47
Metal	3.58
Metal working	3.38
Machinery	12.84
Electrical machinery	40.53
Electronics, instruments	19.47
Vehicles	7.81
Watches	4.85
Other manufacturing	2.28
Firm size	
Small (5-49)	4.58
Medium-sized (50-249)	6.92
Large (> 250)	55.76

Table 2.A.6: Basic Model: Comparison of Two Different Measures of Knowledge Spillovers, GLS Random Effects Estimates

	(1) $LINNS_t$	(2) $LINNS_t$	(3) $LINNS\_N_t$	(4) $LINNS\_M_t$	(5) $LINNS_t$	(6) $LINNS\_N_t$	(7) $LINNS\_M_t$
D_{t-1}	0.064	0.056	0.126**	0.061	0.058	0.127**	0.063
	(0.046)	(0.047)	(0.052)	(0.045)	(0.047)	(0.052)	(0.045)
$NCOMP_{t-1}$	-0.010	-0.012	0.019	-0.003	-0.011	0.020	-0.003
	(0.030)	(0.030)	(0.032)	(0.032)	(0.030)	(0.032)	(0.031)
$LEMPL_{t-1}$	0.964***	0.911***	0.889***	0.061	0.903***	0.881***	0.920***
	(0.039)	(0.042)	(0.045)	(0.045)	(0.042)	(0.046)	(0.050)
$HQUAL_{t-1}$	0.010***	0.009***	0.006*	0.009***	0.009***	0.006*	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$FOREIGN_{t-1}$	0.188**	0.184**	0.223**	0.258***	0.200**	0.232**	0.271***
	(0.093)	(0.093)	(0.100)	(0.095)	(0.091)	(0.099)	(0.096)
$APPR_{t-1}$	0.022	0.018	0.065*	-0.019	0.018	0.065*	-0.018
	(0.032)	(0.032)	(0.035)	(0.036)	(0.032)	(0.035)	(0.036)
LK_{t-1}	0.128***	0.123***	0.182***	0.069	0.129***	0.184***	0.080
	(0.037)	(0.048)	(0.053)	(0.051)	(0.046)	(0.051)	(0.049)
$LSPILL\_ALL_{t-1}$	0.053*						
	(0.030)						
$LSPILL\_CH_{t-1}$		0.099***	0.071**	0.113***			
		(0.028)	(0.031)	(0.029)			
$LSPILL_{t-1}$					0.094***	0.073***	0.098***
					(0.028)	(0.028)	(0.029)
Industry dummies (15)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (4)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Const.	10.448***	10.855***	10.364***	9.904***	10.752***	10.286***	9.765***
	(0.376)	(0.375)	(0.363)	(0.387)	(0.362)	(0.355)	(0.372)
N	701	696	653	628	696	653	628
Wald chi2	1495.97***	1555.93***	1314.47***	1318.99***	1562.36***	1286.87***	1300.46***
R2 overall	0.72	0.73	0.69	0.73	0.73	0.69	0.72
R2 between	0.75	0.75	0.72	0.75	0.75	0.72	0.75
R2 within	0.05	0.05	0.06	0.03	0.05	0.06	0.03
Rho	0.44	0.42	0.46	0.34	0.42	0.46	0.35

*,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 2.A.7: GLS Random Effects Estimates; Competition Effects

	(1) $LINNS_t$	(2) $LINNS\_N_t$	(3) $LINNS\_M_t$	(4) $LINNS_t$	(5) $LINNS\_N_t$	(6) $LINNS\_M_t$
D_{t-1}	0.062	0.132**	0.066	0.049	0.117**	0.062
	(0.047)	(0.052)	(0.045)	(0.047)	(0.052)	(0.045)
$LEMPL_{t-1}$	0.908***	0.887***	0.924***	0.906***	0.883***	0.924***
	(0.042)	(0.046)	(0.050)	(0.042)	(0.046)	(0.050)
$HQUAL_{t-1}$	0.009***	0.006*	0.010***	0.009***	0.006**	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$FOREIGN_{t-1}$	0.196**	0.222**	0.268***	0.224**	0.254**	0.286***
	(0.091)	(0.099)	(0.096)	(0.091)	(0.099)	(0.096)
$APPR_{t-1}$	0.019	0.067*	-0.017	0.023	0.068*	-0.014
	(0.032)	(0.035)	(0.036)	(0.032)	(0.035)	(0.035)
LK_{t-1}	0.125***	0.177***	0.077	0.133***	0.185***	0.087*
	(0.045)	(0.050)	(0.048)	(0.045)	(0.049)	(0.048)
$LSPILL_{t-1}$	0.103***	0.085***	0.103***	0.102***	0.082***	0.098***
	(0.028)	(0.029)	(0.030)	(0.028)	(0.029)	(0.030)
$NCOMP_{t-1}$	0.039	0.088**	0.031			
	(0.037)	(0.045)	(0.044)			
$LSPILL \times NCOMP_{t-1}$	-0.053*	-0.072**	-0.035			
	(0.030)	(0.030)	(0.034)			
$NCOMP > 15_{t-1}$				0.217	0.355*	0.174
				(0.168)	(0.190)	(0.193)
$LSPILL \times NCOMP > 15_{t-1}$				-0.075*	-0.098**	-0.053
				(0.041)	(0.040)	(0.049)
Industry dummies (15)	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (4)	Yes	Yes	Yes	Yes	Yes	Yes
Const.	10.654***	10.156***	9.697***	10.788***	10.399***	9.775***
	(0.359)	(0.361)	(0.373)	(0.353)	(0.346)	(0.361)
N	696	653	628	702	659	632
Wald chi2	1555.2***	1280.05***	1292.04***	1588.58***	1297.93***	1284.20***
R2 overall	0.73	0.70	0.72	0.73	0.69	0.72
R2 between	0.75	0.72	0.75	0.75	0.71	0.75
R2 within	0.05	0.07	0.03	0.06	0.07	0.03
Rho	0.42	0.46	0.35	0.40	0.45	0.35

*,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 2.A.8: GLS Random Effects Estimates; Regional Effects

	(1) $LINNS_t$	(2) $LINNS_t$	(3) $LINNS_t$	(4) $LINNS_t$	(5) $LINNS_t$
D_{t-1}	0.057	0.056	0.058	0.062	0.063
	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)
$NCOMP_{t-1}$	-0.010	-0.012	-0.010	-0.010	-0.010
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
$LEMPL_{t-1}$	0.916***	0.911***	0.933***	0.951***	0.960***
	(0.043)	(0.042)	(0.042)	(0.043)	(0.041)
$HQUAL_{t-1}$	0.008***	0.009***	0.009***	0.010***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$FOREIGN_{t-1}$	0.171*	0.184**	0.171*	0.195**	0.195**
	(0.093)	(0.093)	(0.093)	(0.094)	(0.094)
$APPR_{t-1}$	0.021	0.018	0.021	0.019	0.021
	(0.032)	(0.032)	(0.032)	(0.032)	(0.033)
LK_{t-1}	0.126**	0.123***	0.129***	0.178***	0.192***
	(0.053)	(0.048)	(0.049)	(0.047)	(0.049)
$LSPILL\_CH_{t-1}$	0.088***	0.099***			
	(0.031)	(0.028)			
$LSPILL\_EU_{t-1}$	0.028		0.063***		
	(0.025)		(0.024)		
$LSPILL\_US_{t-1}$	-0.018			0.016	
	(0.027)			(0.026)	
$LSPILL\_JP_{t-1}$	-0.021				-0.006
	(0.030)				(0.027)
Industry dummies (15)	Yes	Yes	Yes	Yes	Yes
Year dummies (4)	Yes	Yes	Yes	Yes	Yes
Const.	10.835***	10.855***	10.771***	10.588***	10.524***
	(0.388)	(0.375)	(0.387)	(0.392)	(0.377)
N	696	696	696	696	696
Wald chi2	1584.50***	1555.93***	1515.15***	1531.74***	1571.40***
R2 overall	0.73	0.73	0.73	0.73	0.73
R2 between	0.75	0.75	0.75	0.75	0.75
R2 within	0.05	0.05	0.05	0.05	0.05
Rho	0.42	0.42	0.43	0.43	0.43

*,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

2.B Further Tables

Table 2.B.1: Correlations of the Model Variables

	1	2	3	4	5	7	8	9	10	11	12	13	14	15	16	17	18
1 $LINNS_t$	1.000																
2 $LINNS\_N_t$	0.9458	1															
3 $LINNS\_M_t$	0.955	0.8247	1														
4 D_{t-1}	0.030	0.0882	0.0548	1.000													
5 $NCOMP_{t-1}$	-0.002	0.0462	0.0205	0.006	1.000												
6 $LEMPL_{t-1}$	0.814	0.8041	0.8179	-0.031	0.034												
7 $HQUAL_{t-1}$	0.184	0.1693	0.1853	0.084	-0.117	1.000											
8 $FOREIGN_{t-1}$	0.058	0.0653	0.0733	0.027	-0.086	0.071	1.000										
9 $APPR_{t-1}$	0.073	0.0974	0.0598	-0.008	-0.007	-0.013	0.099	1.000									
10 LK_{t-1}	0.569	0.5745	0.5425	0.031	-0.057	0.167	-0.041	-0.002	1.000								
11 $LSPILL_{t-1}$	0.557	0.5549	0.5381	0.101	-0.008	0.265	0.000	0.050	0.584	1.000							
12 $LSPILL_CH_{t-1}$	0.557	0.5539	0.5464	0.128	0.006	0.252	0.020	0.038	0.592	0.804	1.000						
13 $LSPILL_EU_{t-1}$	0.511	0.5328	0.516	0.112	-0.013	0.254	0.068	-0.005	0.616	0.662	0.757	1.000					
14 $LSPILL_US_{t-1}$	0.461	0.4836	0.4649	0.114	0.004	0.171	0.006	0.045	0.546	0.489	0.593	0.572	1.000				
15 $LSPILL_JP_{t-1}$	0.351	0.3749	0.3651	0.088	0.014	0.139	-0.089	0.057	0.520	0.340	0.431	0.432	0.420	1.000			
16 $LSPILL_APP_{t-1}$	0.557	0.5578	0.533	0.120	0.000	0.286	0.005	0.058	0.579	0.982	0.812	0.666	0.498	0.336	1.000		
17 $LSPILL_FIRMS_{t-1}$	0.556	0.5598	0.5355	0.093	-0.003	0.269	-0.004	0.049	0.568	0.971	0.789	0.656	0.485	0.334	0.970	1.000	
18 $LSPILL_BACK_{t-1}$	0.378	0.3858	0.3856	0.064	0.020	0.253	0.003	0.039	0.309	0.790	0.587	0.401	0.313	0.244	0.765	0.760	1.000

Table 2.B.2: GLS Random Effects

	(1) $LINNS_t$	(2) $LINNS_t$	(3) $LINNS_t$	(4) $LINNS_t$	(5) $LINNS_t$	(6) $LINNS_t$
D_{t-1}	0.055	0.059	0.056	0.063	0.065	0.065
	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)
$NCOMP_{t-1}$	-0.012	0.039	-0.010	0.040	-0.011	0.007
	(0.030)	(0.037)	(0.030)	(0.037)	(0.030)	(0.034)
$LEMPL_{t-1}$	0.902***	0.907***	0.935***	0.912***	0.932***	0.934***
	(0.042)	(0.042)	(0.040)	(0.043)	(0.040)	(0.040)
$HQUAL_{t-1}$	0.009***	0.009***	0.009***	0.009***	0.009***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
$FOREIGN_{t-1}$	0.195**	0.192**	0.191**	0.199**	0.201**	0.199**
	(0.092)	(0.091)	(0.094)	(0.091)	(0.092)	(0.091)
$APPR_{t-1}$	0.018	0.019	0.015	0.019	0.020	0.021
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
LK_{t-1}	0.129***	0.123***	0.130***	0.131***	0.180***	0.178***
	(0.046)	(0.045)	(0.050)	(0.046)	(0.045)	(0.045)
$LSPILL\_APP_{t-1}$	0.093***	0.102***				
	(0.027)	(0.028)				
$LSPILL\_APP * NCOMP_{t-1}$		-0.055*				
		(0.030)				
$LSPILL\_FIRMS_{t-1}$			0.107**	0.094***		
			(0.044)	(0.028)		
$LSPILL\_FIRMS * NCOMP_{t-1}$				-0.056*		
				(0.031)		
$LSPILL\_BACK_{t-1}$					0.141***	0.149***
					(0.054)	(0.057)
$LSPILL\_BACK * NCOMP_{t-1}$						-0.075
						(0.105)
Industry dummies (15)	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies (4)	Yes	Yes	Yes	Yes	Yes	Yes
Const.	10.811***	10.770***	10.778***	10.666***	10.582***	10.549***
	(0.368)	(0.370)	(0.388)	(0.381)	(0.363)	(0.363)
N	696	696	696	696	696	696
Wald chi2	1564.275***	1554.801***	1546.643***	1540.510***	1538.597***	1541.555***
F-statistic						
R2 overall	0.731	0.733	0.728	0.732	0.728	0.728
R2 between	0.752	0.753	0.748	0.752	0.751	0.751
R2 within	0.048	0.052	0.050	0.052	0.047	0.048
Rho	0.426	0.426	0.430	0.422	0.425	0.425

*,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 2.B.3: OLS Estimates

	(1)	(2)
	LINNSt	LINNSt
$LINNS_{t-1}$	0.408***	0.409***
	(0.049)	(0.049)
D_{t-1}	0.048	0.052
	(0.045)	(0.045)
$NCOMP_{t-1}$	0.023	0.079*
	(0.029)	(0.041)
$LEMPL_{t-1}$	0.492***	0.497***
	(0.059)	(0.059)
$HQUAL_{t-1}$	0.007***	0.007***
	(0.002)	(0.002)
$FOREIGN_{t-1}$	0.115	0.106
	(0.082)	(0.081)
$APPR_{t-1}$	0.031	0.031
	(0.032)	(0.032)
LK_{t-1}	0.096**	0.091**
	(0.042)	(0.042)
$LSPILL_{t-1}$	0.074***	0.082***
	(0.025)	(0.025)
$LSPILL \times NCOMP_{t-1}$		-0.058*
		(0.030)
Industry dummies (15)	Yes	Yes
Year dummies (4)	No	No
Const.	6.446***	6.308***
	(0.604)	(0.601)
N	640	640
F-statistic	96.29***	93.04***
R2	0.76	0.77

*,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 2.B.4: Blundell and Bond Estimates

	(1) $LINNS_t$	(2) $LINNS_t$	(3) $LINNS_t$
$LINNS_{t-1}$	-0.014	0.020	-0.022
	(0.127)	(0.115)	(0.110)
D_t	0.089	0.084	0.099
	(0.060)	(0.067)	(0.063)
$NCOMP_t$	-0.046	-0.067	-0.040
	(0.061)	(0.064)	(0.059)
$LEMPL_t$	0.928***	0.804***	0.882***
	(0.200)	(0.201)	(0.159)
$HQUAL_t$	0.006	0.006	0.006
	(0.004)	(0.004)	(0.004)
$FOREIGN_t$	-0.008	0.036	0.019
	(0.168)	(0.183)	(0.147)
$APPR_t$	0.019	0.022	0.017
	(0.044)	(0.047)	(0.047)
LK_t	-0.020	-0.023	0.086
	(0.161)	(0.163)	(0.119)
$LSPILL_t$	0.056**	0.065***	0.059***
	(0.024)	(0.025)	(0.022)
Const.	11.678***	11.752***	11.730***
	(2.404)	(2.142)	(1.626)
N	721	721	721
Wald chi2	37.008***	40.783***	162.102***
N instruments	23	30	40
Sargan test (chi2)	16.76	24.56	30
Sargan test (p-value)	0.080	0.219	0.451
AR(1) p-value	0.003	0.002	0.004
AR (2) p-value	0.771	0.818	0.718

*,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

3 Technological Diversity, Uncertainty, and Innovation Performance*

3.1 Introduction

The main goal of this paper is to achieve a better understanding of the effect of technological diversity on innovation in different stages of a firm's innovation process and how these relationships are affected when a firm faces technological uncertainty. There is a comprehensive empirical literature investigating the relationship between technological diversity and the number of patent applications, an intermediate innovation output. Such studies proclaim a positive impact of diversity on the number of patent applications (see, e.g., Garcia-Vega 2006; Leten et al. 2007). Furthermore, a number of studies analyses the impact of diversity on firm performance in terms of profitability (e.g., Chiu et al. 2008; Kim et al. 2009; Miller 2006) or corporate growth (Granstrand 1998; Kim et al. 2016).

While the literature analyzes the effect of technological diversity on innovation input and financial firm performance, little knowledge exists how diversity affects the commercial success of innovative products. The most closely related literature from the pharmaceutical industry suggests that the relevance of technological diversity differs between the processes of discovering a molecule and obtaining application approval through clinical trials (see, e.g., Arora et al. 2009; Cockburn and Henderson 2001; Danzon et al. 2005; Henderson and Cockburn 1996; Plotnikova et al. 2010). However, the process of clinical trials differs starkly from the product development process in other industries, particularly because patents of pharmaceutical molecules often represent final products, while most final products in other industries require a

*This chapter is co-authored with Thomas Bolli and Martin Wörter substantive development process of the underlying technology. Furthermore, the existing empirical evidence from the pharmaceutical industry does not account for potential endogeneity of technological diversity.

This paper aims to fill this research gap by analysing the relationship between technological diversity and innovation performance for a broader sample of manufacturing firms using panel data estimates. It investigates whether the role of technological diversity differs between innovation input, patent success and innovation output of a firm as measured by the sales share of innovative products. Our data improves upon existing estimates that analyse the impact of diversity on a binary outcome for successful market introduction by evaluating the relationship between diversity and the quantitative product market success, i.e., the share of sales generated by new and improved products.

We show that the findings at the technological level (patents) do not necessarily apply at the level of product markets. Our results confirm the established findings that technological diversity increases the number of patent applications. In contrast, our analysis of innovation output, suggests a negative impact of diversity on the sales share of innovative products.

Although the paper's main contributions are empirical, we draw on an organizational learning framework in order to conceptualize the potential tension between discovery and commercialisation that allows for a differing impact of diversity also from a conceptual point of view. Referring to the distinction between exploration and exploitation activities (March 1991), we argue that exploration activities are emphasized in the discovery stage (technological development) and exploitation activities are more important in later stages of the innovation process (commercialization). Given this distinction, we can expect that technological diversity is more important for explorative competences and much less important for exploitative ones. High coordination costs in later stages of the development process and cognitive tensions between organizational units representing different stages of the innovation process suggest different impacts of technological diversity compared to the discovery stage.

We also address the moderating role of technological uncertainty for the relationship between technological diversity and innovation performance both conceptually and empirically. Technological uncertainty increases the costs of technological development and may alter the importance of technological diversity in the innovation process. We argue that technological uncertainty moderates the impact of technological diversity by further deepening the gap between the discovery and commercialization stage. Firms then need to rebalance their activities associated with both stages that increases the costs in both of them. Therefore, the increase in costs driven by uncertainty is supposed to simultaneously decrease the benefits of technological diversity. Our empirical investigation indeed not only shows that technological diversity increases patenting and decreases the sales share with new products, but that technological uncertainty negatively moderates the effect of diversity on innovation for outcomes in both stages, namely the number of patents and the sales share.

Our study is based on a comprehensive data set combining firm-level panel data stemming from the Swiss Innovation Survey (equivalent to the European 'Community Innovation Survey') with information about patent applications and the corresponding International Patent Classification (IPC) inscriptions for the panel firms. The resulting data set allows us to compare the impact of technological diversity on alternative measures of innovation performance controlling for other important innovation drivers. We contrast the traditional measures 'research and development (R&D) intensity' and 'number of patent applications' to capture the research output of the discovery stage while the widely applied innovation measure 'share of innovative sales' approximates the commercialization stage (see, e.g., Belderbos et al. 2004; Miotti and Sachwald 2003).

The paper is organized as follows. Section 3.2 discusses the relationship between exploration-exploitation, technological diversity and innovation as well as the moderating role of technological uncertainty. This discussion leads to our research hypotheses. Section 3.3 presents the employed data set and econometric specification. Section 3.4 discusses the results and Section 3.5 concludes the paper.

3.2 Literature Review and Hypotheses

3.2.1 Technological Diversity, Product Diversity and Strategic Diversification

Technological diversity might be part of a corporate diversification strategy, but corporate diversification comprises many business aspects and most of them, such as diversification of business segments, are not technology-related. Torrisi and Grandstrand (2004) designate diversification as a (strategy-driven) process by which the range or diversity of a product or technology portfolio is increased by adding elements of a new type. In this paper, we refer to technological diversity in order to describe the breadth of the technological portfolio of a firm (Granstrand et al. 1997; Miller 2006) and abstain from the question whether it is a result of an overarching corporate strategy or not.

According to Granstrand (1998), many high-growth firms seem to follow a two-stage process starting with a process to increase technological diversity before increasing product diversity. This means that technological diversity can be a necessary condition for product diversity. In addition, technological diversity is very often larger than product diversity so that firms have a set of new technologies at hand they can choose from, recombine them in unique ways and exploit them for new product developments (Pavitt 1998). The accumulation of capabilities in technological fields outside a firm's core technologies allows a firm to monitor and absorb new knowledge, thereby mitigating potential disruptive effects from new technologies (perhaps introduced by rivals). This process also allows multi-technology corporations to combine new technologies with old technical capabilities in order to follow new 'technological trajectories'. In addition to exploiting arising technological opportunities, investing in multiple technological areas allows for the exploitation of economies of scope and spillovers (Torrisi and Grandstrand 2004). Economies of scope and spillovers are underlying factors of what Breschi et al. (2003) call knowledge relatedness from an empirical point of view. In this context, they find evidence that firms mainly diversify their innovative activities across related technological fields, i.e., fields that share a common knowledge base.

3.2.2 The Relationship between Technological Diversity and Innovation Performance in Different Innovation Stages

Exploration and Exploitation in the Discovery and Commercialization Stage

In order to derive our hypotheses regarding the impact of diversity in the discovery and commercialization stage of the innovation process, we build upon the theoretical framework of exploration and exploitation (March 1991), where exploration refers to the development of new knowledge and exploitation refers to using already existing knowledge (Levinthal and March 1993).

Although March initially referred to a trade-off between exploration and exploitation, it is common sense that the trade-off is not a zero-sum game; actually, both activities are complementary and mutually enhancing and need to be balanced within an organization (see, e.g., Chen and

Katila 2008; Lavie et al. 2010).[1] Explorative and exploitative activities fulfil different functions that are both essential elements of the innovation process. In order to put them into effect one has also to look at the organizational dimension, i.e., how the functions and related tasks are assigned to organizational units. An organization has to ensure that it can pursue both explorative and exploitative activities to a sufficient degree by assigning the related tasks to corresponding organizational units.[2] In particular, to achieve higher innovation performance, a firm has to co-specialize in exploration and exploitation across separate sub-units (Simsek 2009), calling for organizational consequences.

In order to relate exploration and exploitation to the discovery and commercialization stage of the innovation process, we refer to the function domain perspective (Li et al. 2008).[3] This approach relates exploration to technology search and exploitation to product market knowledge search (see Danneels 2002; Nerkar and Roberts 2004), suggesting that upstream activities are more explorative than activities that are located downstream in the value chain. Hence, the functional domain approach of the exploration-exploitation framework suggests that exploration is relatively more relevant in the discovery stage, while exploitation is relatively more relevant in the commercialization stage. The corresponding organizational units that are associated with different stages of the innovation process immediately follow from the main functions they have to fulfil: The R&D department is especially relevant in the discovery stage, whereas the product development and marketing department is more relevant in commercialization stage. Especially high-tech companies are supposed to simultaneously engage in a high degree of exploration in R&D and a high degree of exploitation in complementary domains such as manufacturing, sales, and service (Gupta et al. 2006). The literature analyzing different types of alliances that support either exploration or exploitation activities supports this view (Hagedoorn and Duysters 2002; Lavie and Rosenkopf 2006). Generally, this literature argues that alliances with a focus on R&D aim to explore technological opportunities and hence represent exploitative alliances (Rothaermel and Deeds 2004). Conversely, alliances that focus on manufacturing, marketing

[1] For empirical evidence that balancing enhances performance measures, see, e.g., He and Wong (2004); Hill and Birkinshaw (2006); Katila and Ahuja (2002). He and Wong (2004) find that the interaction between explorative and exploitative innovation strategies is positively related to the sales growth rate, whereas the relative imbalance between them is negatively related to it.

[2] Generally, one can distinguish between integration of exploitative/explorative tasks within the same unit and differentiation of tasks across units, i.e., the subdivision into distinct organizational units (see Blindenbach-Driessen and Ende 2014; Lavie et al. 2010). According to Raisch et al. (2009), they are complementary mechanisms and a subdivision is not always clear-cut.

[3] The knowledge distance domain represents another dimension (see Li et al. 2008) where a firm can search for knowledge that is either local or distant to the existing knowledge stock within each function domain.

or supply agreements aim to exploit market opportunities and hence represent exploitative alliances (Rothaermel 2001).

An essential question is whether exploitation depends upon prior exploration, suggesting a sequential process where the R&D department first explores new technologies or new combinations before the marketing department exploits market opportunities. This might be a good description of discrete-product or science-driven industries such as the pharmaceutical industry and it also follows the sequential logic of technological and product diversity depicted above, but in many industries, the relationship might be more complex and such a linear process highly stylized. First, firms in more complex industries might switch back and forth between units that are more exploratory and those that are more exploitative. Second, units representing the commercialization stage can also engage in explorative activities, especially if one thinks about exploring knowledge from external sources.[4] However, for the analysis of the impact of technological diversity in technology-driven firms, it is fair to distinguish conceptually between a discovery stage where the R&D department is mainly involved in exploration of (diversified) technologies and a commercialization stage where the marketing and sales department is mainly involved in exploitation of discoveries by commercializing them.[5]

In sum, we contend that firms that engage in exploitation and exploration at the same time install organizational units that pursue both exploration and exploitation, but to different degrees, i.e., there are units that are more oriented towards exploration while there are also units that are more oriented towards exploitation. They represent different stages of the innovation process and have different value chain functions where upstream activities are more exploratory than downstream activities. Therefore, in line with the function domain approach, we suppose that technology-driven firms balance exploration and exploitation across units involving R&D labs to a higher degree in explorative tasks and the sales and marketing departments to a higher degree in more standardized, exploitative tasks.

[4] 'Open innovation' (Chesbrough 2006) and 'user innovation' (Von Hippel 2005) are prominent features often discussed in literature with the goal of opening up also later stages of the innovation process and to explore from external knowledge sources.

[5] In the discovery stage, we can only capture the process that leads to a patented invention empirically. However, we can claim that our results at least hold for more technologically-driven discoveries leading to patents.

The Effect of Diversity on Explorative and Exploitative Innovation Performance

The interplay between organizational units has been studied extensively and especially the R&D-marketing relationship is associated with organizational tensions and high coordination costs that are due to different cultures, attitudes and organizational and communication practices (see, e.g., Gupta et al. 1986; Ruekert and Walker Jr 1987; Souder 1981). A gap between R&D and marketing might lead to substantial different effects of technological diversity in different development stages as the marketing department might struggle to apply technologically complex discoveries from the R&D department. As Granstrand et al. (1997) put it, 'failure to exploit radically new technologies has more to do with failure in product development, production, marketing, and organizational adaptation than with failure in technological competencies'. Interestingly, although the cognitive gap between organizational units and the resulting risk of failure is recognized, most of the literature presumes that technological diversity is associated with better firm performance and increases the ability to appropriate returns from R&D activities (Torrisi and Grandstrand 2004).

A cognitive gap between different organizational units related to increasing technological diversity also increases coordination and communication costs (Quintana-García and Benavides-Velasco 2008) and consequently reduces the benefits of technological diversity for downstream exploitative activities. Furthermore, according to Quintana-García and Benavides-Velasco (2008), the benefits of technological diversity in terms of economies of scope and spillovers are larger for explorative activities than for exploitative activities. Hence, they suggest that technological diversity improves the performance of explorative activities more than the performance of exploitative activities.

The Effect of Diversity on Innovation Performance in the Discovery and Commercialization Stage

Referring to the organizational representations of innovation activities and the inherent R&D-marketing tension, we extend Quintana-Garcia and Benavides-Velasco's argument to different stages of the innovation process and the organizational units that represent these stages, namely the R&D department in the discovery stage and the product development and marketing department in the commercialization stage.[6] Consequently, we argue that technological diversity should increase innovation performance in the discovery stage more than in the commercialization stage. We hypothesize technological diversity will increase the innovation input and intermediate output in the discovery stage, namely R&D intensity and the number of patent applications. Conversely, we expect it to decrease commercial success of these discoveries. The latter is due to the mainly exploitative character of this form of commercialization and the R&D-marketing tension described above involving high coordination costs especially in later stages of the innovation process.

Hence, combining the function domain approach of the exploration-exploitation framework with the argument of Quintana-Garcia and Benavides-Velasco (2008) that technological diversity increases explorative innovation performance more than exploitative innovation performance, we hypothesize that

Hypothesis 1: Technological diversity increases innovation in the discovery stage.

Hypothesis 1a: Technological diversity increases R&D intensity.

Hypothesis 1b: Technological diversity increases the number of patent applications.

Hypothesis 2: Technological diversity decreases innovation performance in the commercialization.

Hypothesis 2a: Technological diversity decreases the sales share of innovative products.

3.2.3 Technological Uncertainty and Coordination Costs

Technological and market uncertainty is created by dynamic environments that may be characterized by changes in technologies, variations in customer preferences, and fluctuations in product demand or supply of materials (Fleming 2001; Jansen et al. 2006). In turbulent environments, there is a need to make risky investments as it becomes more important to bring products to market in a timely manner (Calantone et al. 2003).[7] Following Granstrand (1998),

[6] In contrast to Quintana-García and Benavides-Velasco (2008), we treat exploration and exploitation as innovation activities that are part of the innovation process rather than outcomes (see Li et al. 2008). Consequently, we do not explicitly measure them.

[7] Although not explicitly discussed in the literature, technological uncertainty might also arise from overly complex technological dynamics. When it comes to commercialization, technological complexity is often identified as a major impediment (Singh 1997). Singh argues that firms developing technologically more complex technologies face greater difficulties in developing the required competencies and greater organizational costs than those

the (uncertain) environment that creates both threats and opportunities must be recognized explicitly in the analysis of technological diversity.

In general, organizations have higher information processing needs in uncertain environments than in normal situations (Gupta et al. 1986). The rise of coordination costs can be localized at the interplay between organizational units (i.e., between R&D and marketing) (Bstieler 2005; Calantone et al. 2003; Calantone and Rubera 2012). Uncertainty even increases the disharmony between R&D and marketing so that additional integration measures are necessary and technology management appears to be increasingly important in order to deal with environmental challenges (Granstrand 1998). Under uncertainty, R&D and marketing need to exchange information more frequently to keep pace with technological and market changes (Ruekert and Walker Jr 1987), collaboration between R&D and marketing is assumed to be more important (Gupta et al. 1986), and cross-functional integration is essential (Eng and Ozdemir 2014; Song and Thieme 2006). Firms therefore have to constantly rebalance exploration with exploitation activities to remain competitive. This creates a need for organizational consequences across units (Chen and Katila 2008) that increases the costs in both the discovery and the commercialization stage simultaneously. Therefore, the uncertainty driven change in costs is supposed to be constant between exploitation and exploration.

To summarize, we argue that technological uncertainty increases information and coordination costs significantly and influences the effect of diversity on innovation negatively. Further increasing costs of diversity consequently decreases the effect of diversity on R&D intensity resp. patent applications and the sales share of innovative products that are the outcomes in the two stages. Consequently, we formulate the following hypothesis with respect to a moderating effect of technological uncertainty:

Hypothesis 3a: The diversity-innovation effect is negatively moderated by technological uncertainty for both R&D intensity and the number of patent applications (discovery stage) and the sales share of innovative products (commercialization stage).

developing less complex technologies. Consequently, they bear a much larger risk of failure in commercializing.

3.3 Data and Methodology

The employed panel data stems from five waves of the Swiss innovation survey conducted by the KOF in the years $t = \{1996, 1999, 2002, 2005, 2008\}$, where t denotes the time period. The surveys are based on a disproportionately stratified random sample of firms with more than five employees (full time equivalents) covering the most important industries of the manufacturing, construction and service sector. Stratification takes place on industry and within each industry on three firm size classes. Responses were received from 1,989 (32.5%), 2,172 firms (33.8%), 2,583 firms (39.6%), 2,555 firms (38.7%), and 2,363 (36.1%) for the years 1996, 1999, 2002, 2005 and 2008, respectively. However, the investigation at hand only uses data from manufacturing firms. Dropping observations with missing values yields a highly unbalanced firm-panel with 3,110 observations.

We enrich the innovation survey with annual information about patent applications from the European Patent Office (EPO 2013) and the Derwent World Patent Index (WPI) by Thomson Reuters. This allows us to construct the existing patent stock of a firm and the number of new patent applications in a period.[8] Given the three-year periodicity of the innovation survey, we define the number of new patent applications ($New\_Patent\_Applications_{it}$) of firm i in period t as the sum of patent applications over the corresponding three years. Following the perpetual inventory method (see Cockburn and Griliches 1988), the patent stock ($Patent\_Stock_{it}$) of firm i in period t refers to the depreciated sum of patent applications in the six years before the period, where we follow the literature in assuming a geometric discounting process with a depreciation rate of 15% (see, e.g., Aghion et al. 2012; Keller 2002a).

The patent data also entails information about the patent section and class inscription (IPC code) on different levels of aggregation. We use the class level with three digits and the section level with one digit (for further explanations, see WIPO 2014). Patent examiners use IPC codes to categorize a patent application according to the underlying technologies. Therefore, these technological fields entail important information about the technological content of a patent.

We use different measures of technological diversity that have been suggested in literature: First,

[8] We conducted several rounds of names matching: First, we used all patent applicants from Swiss applicants from WPI between 1990 and 2010, cleaned the applicants' names and firm names, and matched the cleaned applicants' names with firm names from the Innovation Survey automatically with a matching software. Afterwards, we checked the results manually. We also searched each firm name from the panel in ESPACENET and PATSTAT to get as many as possible patent applications. At the end, all matched patent applications we found were compiled in one dataset and checked once again. For the analysis here, we use patent families rather than single applications. Families comprise multiple applications in different countries. They are better able to reflect inventions than single patent applications (Martínez 2011; OECD 2009).

we report results based on a simple Herfindahl index and on the entropy measure suggested by Jacquemin and Berry (1979).[9] Concretely, the first measure refers to one minus the Herfindahl index, calculated as the sum of squared patent section inscription shares:

$$Diversity_{it} = 1 - \sum_c \left(\frac{N_{scit}}{N_{it}}\right)^2 \tag{3.3.1}$$

N_{scit} denotes the discounted number of patent section inscriptions of firm i's patent stock in patent section s in patent class c in period t. The discounted patent stock in period t, N_{it}, refers to the discounted sum of patent applications in the six years before the respective period.
The second diversity measure refers to the entropy measure suggested by Jacquemin and Berry (1979):

$$Entropy_{it} = \sum_c \frac{N_{scit}}{N_{it}} \times ln\left(\frac{N_{it}}{N_{scit}}\right) \tag{3.3.2}$$

Following Jacquemin and Berry (1979), we distinguish between related and unrelated diversity. The first, related diversity, measures the diversity of patent classes within the patent sections. N_{sit} represents the sum of the discounted number of patent section inscriptions of firm i's patent stock in patent section s in period t across patent classes c.

$$Entropy_{it} = \sum_{c \in s} \frac{N_{scit}}{N_{sit}} \times ln\left(\frac{N_{sit}}{N_{scit}}\right) \tag{3.3.3}$$

Conversely, unrelated diversity captures diversity across technological sections, which is given by

$$Entropy_{it} = \sum_s \frac{N_{sit}}{N_{it}} \times ln\left(\frac{N_{it}}{N_{sit}}\right) \tag{3.3.4}$$

where N_{it} refers to the discounted sum of patent applications in the six years before the period. This paper employs mainly three types of dependent variables that cover the discovery and the subsequent commercialization stage (see Table 3.A.1). Following the existing literature (see, e.g., Garcia-Vega 2006; Leten et al. 2007), the first type refers to innovation inputs, namely R&D intensity ($R\&D\_Intensity_{it}$). The second type refers to intermediate outputs, namely

[9] For a comprehensive overview of diversity measures, see Dawson (2012); Stirling (1998).

patent applications in the current three-year period, t, ($New\_Patent\_Applications_{it}$). The third type covers the commercialization of innovation on the market and is the sales share with innovative products. Following Garcia-Vega (2006); Leten et al. (2007), we control for R&D intensity ($R\&D\_Intensity_{it}$), firm size ($Size_{it}$) and the existing patent stock ($Patent\_Stock_{it}$). The diversity can take the value 0 either because a firm has no patents or because all patents fall into a single patent section. Hence, we further include a dummy variable indicating whether the patent stock of a firm is empty (Pat_dummy_{it}) in addition to a variable that captures whether a patent stock is based on a single patent ($Single\_pat_{it}$). Thereby, we control for the large number of firms with a single patent. These firms necessarily have a diversity of zero though this might not necessarily reflect a specialization choice. In order to capture the differences in the ability of patents in protecting innovations, we further include a dummy variable that indicates whether protection measures (e.g. patents, copyrights, secrecy) are effective ($Protection_{it}$). Patents and other innovation outputs might be an outcome from inventor collaborations across firms or might depend on exploration from external knowledge sources. Hence, we also control for incoming knowledge spillovers from customers ($Incoming\_Spillovers\_Customers_{it}$), suppliers ($Incoming\_Spillovers\_Suppliers_{it}$) and competitors ($Incoming\_Spillovers\_Competitors_{it}$) as well as for the presence of R&D collaborations. Measures for price ($Price\_Comp_{it}$) and non-price competition ($Non\_Price\_Comp_{it}$) capture the influence of competition on innovation performance. In addition, we include the share of personnel with tertiary education ($Qualification$) to account for the firm's absorptive capacity. Technological Potential captures the technological potential outside of the firm. Year dummies (α_t) capture unobserved heterogeneity across time (for the summary statistics see Table 3.A.2).

Building the diversity index based on the lagged patent stock in $(t-1)$ and $(t-2)$ accounts for reverse causality. In order to address the problem of unobserved heterogeneity, we include individual intercepts (α_i), i.e., present OLS fixed effects estimates. For the discovery stage, we write our estimation function (with robust standard errors clustered at firm level) for R&D intensity as

$$\begin{aligned} R\&D_{it} = \alpha + \alpha_t + \alpha_i + \beta Diversity_{it} + \gamma_1 Pat\_dummy_{it} \\ + \gamma_2 Single\_patent_{it} + \gamma_3 Patent\_stock_{it} + \delta X_{it} + \epsilon_{1it} \end{aligned} \tag{3.3.5}$$

where X_{it} denotes a vector of control variables for size, appropriability, incoming knowledge spillovers, competition, technological potential and human capital.

The estimation strategy for new patent applications is essentially the same as for R&D intensity. However, we use fixed effects Poisson Pseudo-Maximum-Likelihood models to account for the count data nature of new patent applications. Standard errors are robust, thereby accounting for clustering at the firm level (see, e.g., Cameron and Trivedi 2013). In addition, we include R&D intensity in the patent application equation, i.e., assume that current R&D intensity serves as an input in the production process of the intermediate innovation output, patent applications. Hence, the equation for new patent applications is given by

$$\begin{aligned} Pat\_new_{it} &= \alpha + \alpha_t + \alpha_i + \beta Diversity_{it} + \gamma_1 Pat\_dummy_{it} \\ &+ \gamma_2 Single\_patent_{it} + \gamma_3 Patent\_stock_{it} + \delta_1 R\&D_{it} + \delta_2 X_{it} + \epsilon_{2it} \end{aligned} \tag{3.3.6}$$

The second type of dependent variable refers to the commercialization stage and measures the sales share of innovative products, a widely applied measure of innovation output (see, e.g., Belderbos et al. 2004; Miotti and Sachwald 2003). The OLS estimations with robust standard errors clustered at firm level for the share of sales generated by innovative products entail the same set of control variables as the equation for patent applications.[10] We also estimate this model controlling for unobserved time-invariant heterogeneity with firm-level fixed effects:

$$\begin{aligned} Sales\_share_{it} &= \alpha + \alpha_t + \alpha_i + \beta Diversity_{it} + \gamma_1 Pat\_dummy_{it} \\ &+ \gamma_2 Single\_patent_{it} + \gamma_3 Patent\_stock_{it} + \delta_1 R\&D_{it} + \delta_2 X_{it} + \epsilon_{3it} \end{aligned} \tag{3.3.7}$$

3.4 Results

Table 3.B.1 presents cross-correlations of the variables. As expected, we see a strong positive correlation between diversity measures, patent stock, and new patent applications and between the different measures for diversity (1-Herfindahl, entropy, related entropy, unrelated entropy). This correlation is clearly stronger as compared to the correlation between 'sales share new products' and the different measures of diversity. In addition, 'new patent applications' and 'size' as well as 'R&D intensity' and 'qualification' are clearly positively correlated.

[10] Less than 1% of model predictions lie outside the possible range of 0 to 100. Hence, we use OLS rather than a general linear model such as fractional logit that would be the preferred model for fractional dependent variables (see Papke and Wooldridge 1996). The reason is that OLS allows us to account for unobserved heterogeneity more easily by including firm fixed effects.

Table 3.A.3 summarizes our main results regarding the impact of three diversity measures on three innovation performance measures, namely R&D intensity, patent applications, and the share of sales generated by innovative products (new or essentially improved products). For each innovation performance measure, the table shows the marginal effects for OLS and Poisson and firm-level fixed effects estimations (FE, FE Poisson), respectively.[11]

Hypothesis 1 suggests that technological diversity increases the number of patent applications, whereas hypothesis 2 suggests that it decreases innovation performance as measured by the sales share of innovative products. While the results provide no evidence that technological diversity increases R&D intensity, they support Hypothesis 1 regarding patent applications that represent the discovery stage, thereby confirming the empirical results in the existing literature (Gambardella and Torrisi 1998; Garcia-Vega 2006; Gemba and Kodama 2001; Leten et al. 2007; Nesta and Saviotti 2005). The positive effect of diversity on new patent applications is mainly observed for the unrelated entropy measure, which captures diversity across patent sections. The positive effect vanishes for the related entropy measure if we control for firm-level fixed effects. This means that if diversity is driven by technological activities in different patent sections (large technological distance), positive effects can be expected. The reason why there is no effect on R&D intensity might be that the decision to engage in R&D activities is already taken in the pre-discovery stage. Once a firm has decided to engage in R&D and the corresponding intensity, diversity comes into play in determining the intermediate output that is represented by patent applications.

Going beyond the discovery stage shows that technological diversity decreases the share of sales generated by innovative products as proposed by hypothesis 2. Although the OLS estimations already indicate a negative relationship, we get significant marginal effects if we account for firm-level fixed effects. However, this result is mainly driven by related entropy, i.e., diversity within a technological section. This means that technologically highly specialized firms are likely to have (on average) a significantly higher innovation performance.[12]

The positive diversity effect for patent applications and the negative diversity effect for the innovation equation suggests that diversity is an important driving force for the discovery stage of the innovation process but a hindering force for the commercialization stage of the innovation

[11] Tables 3.B.2-3.B.4 show the estimates containing all control variables.

[12] In estimations that are not shown, we distinguish between the sales share of products that are new and the sales share of products that are only modified, indicating more incremental innovations. The results show that both kinds of innovation outcomes are negatively affected by related entropy.

process. To express it differently, technological specialization does not lead to more patents, but newly developed or essentially improved products based on specialized technologies tend to have a greater market success.

Having in mind that we control for fixed effects, appropriability, the technological profile of a firm, its size and patent activities in the past, etc., it is unlikely that the results are driven by the greater affinity of technologically specialized firms for other measures of appropriability like secrecy or the characteristics of the patent system. Rather, it is likely that an important part of the costs of technological diversity becomes obvious on the market place; diversity results in a greater technological output based on exploratory activities but the departments involved in subsequent development stages with more exploitative activities might be overstrained by the technological complexity. In particular, this is the case, if technological discovery, product development, and commercialization are not carefully balanced and aligned with each other. They might not be able to combine technologically diverse elements into commercially successful products. Therefore, market value of products resulting from this technological output tends to be lower.

If we go one-step further and investigate the moderating effect of technological uncertainty on the relationship between diversity and the performance in the discovery (R&D, patenting) stage and the commercialization stage (innovative sales), respectively, we see that the interaction effect turns out to be negative for both the patent stage and the commercialization stage. Its effect on R&D intensity is insignificant. These results confirm our third hypothesis on the negative moderating effect of technological uncertainty on intermediate and final innovation outcomes. The results hold for the 1-Herfindahl index measure and for the entropy measure (see Table 3.A.4 and Tables 3.B.5-3.B.7). In line with our conceptual framework, these results suggest that uncertainty increases costs for exploration and exploitation simultaneously that mainly take place in the discovery and commercialization stage, respectively. The necessity to balance both activities within an organization that becomes increasingly difficult under technological uncertainty might be the driving force behind our empirical observation.

In addition, we observe positive direct effects of uncertainty on the number of patent applications, and, interestingly on the sales share with innovative products. This suggests that under uncertainty, firms increase both their explorative and exploitative activities, and supports the notion that both activities need to be balanced, i.e., a firm has to put more effort into commercialization when more discoveries are made. Translated to organizational representations of innovation activities, this calls for a careful alignment of R&D and marketing in order to exploit the discoveries successfully. However, in combination with technological diversity, these efforts are undermined as the firms are not able to cope with the inherent complexities arising from both diversity and uncertainty. The interaction between the two simultaneously diminishes both the incentives to make more discoveries and the ability to commercialize the subsequent products on the marketplace successfully.

Part of the literature argues that in dynamic environments, where demand changes rapidly and often in unpredictable ways, exploration is more effective (Jansen et al. 2006) and firms must continuously explore for new opportunities and new technologies (Uotila et al. 2009, p. 222). In contrast, according to Calantone and Rubera (2012), firms tend to overcome the challenge of environment uncertainty by creating specialization. For example, Toh and Kim (2013) consider the scenario where a firm becomes more specialized under technological uncertainty and ambiguity over rivals' actions in order to advance its technologies against competitors. Our results show, in principle, that under technological uncertainty firms would be better off with specialization as we provide empirical evidence for a negative effect of uncertainty on the relationship between diversity and the number of discoveries and commercial success.

As already mentioned in the introduction, the pharmaceutical industry shows some particular characteristics and conceptual frameworks that fit with this industry might not be applicable to other industries with a more complex development process. In addition, from an empirical point of view, it is an industry with a large number of patents, employees etc. and pharmaceutical firms account for about 8% of the sample. Therefore, we conduct two robustness checks (not shown) in order to rule out the possibility that results are driven by this industry and to make sure that our conceptual framework is applicable to other innovation processes as well: First, we estimate the same models excluding the pharmaceutical industry. The results are robust. Second, we estimate the models only for the pharmaceutical industry leaving only about 200 observations. Therefore, some effects become insignificant, but the interaction effect between diversity and uncertainty is still significantly negative.

3.5 Conclusions

Based on comprehensive firm-level panel data comprising a time span of 15 years, we investigate, on the one hand, the relationship between technological diversity and explorative innovation

activities of firms in the so-called discovery stage and, on the other hand, the relationship between technological diversity and the sales share of innovative products that represents the commercialization stage. Our empirical results show that the impact of diversity differs between intermediate innovation output from more explorative activities and innovation output from more exploitative activities, thereby questioning the existing literature that collapses the innovation process into a single stage. Concretely, while we do not find any impact of diversity on R&D intensity, we find a significant and positive relationship between diversity and the number of patent applications. However, we also find a negative direct impact of diversity on the sales share with innovative products. Hence, the results for the patent (discovery) stage in the innovation process cannot be extended to the commercialization stage of innovative products.
In a further set of estimations, we test the moderating effect of technological uncertainty on the relationship between technological diversity and innovation performance variables. The interaction between technological uncertainty and diversity shows a significant and negative effect for new patent applications and the sales share of innovative products.
The first set of results supports the notion that innovation outcomes from activities in different stages of the innovation process are affected differently by technological diversity. Technological diversity increases the success of explorative R&D-related activities and makes subsequent exploitation more costly, which results in the observed negative effect of diversity in the commercialization stage. Organizations need to balance both types of activities across different units in order to remain competitive.
The balancing activity gets even more difficult if the technological environment is uncertain as firms then struggle to align their R&D and marketing activities and the gap between the corresponding organizational units increases. Firms have higher incentives to increase their explorative activities in such circumstances, which unbalances the exploration-exploitation relationship resp. sharpens the R&D-marketing tensions and requires rebalancing and reorganization. Rebalancing increases the costs in both stages what essentially explains why we observe a negative and significant effect of technological uncertainty on the diversity-innovation relationship independent from the stage of innovation process.

Implications for Practice

Our study has a number of implications for managerial practice. First, although we argued that the firms might try balance exploration and exploitation by assigning more explorative tasks to the discovery stage and exploitative tasks to the commercialization stage, our results suggest frictions between the organizational units representing these activities. These frictions lead to the undesirable result that diversity decreases innovation performance on the market. The first implication is therefore to overcome the tension that might be located at the interface between R&D and marketing and might be caused by different communication cultures and attitudes. However, it is very difficult to change cultures and attitudes. Management literature (see, e.g., Griffin and Hauser 1996; Souder 1988) suggests that firms should install regular information exchange mechanisms (e.g., regular meetings, joint workshops, joint other activities) between the units in order to raise mutual understanding and to decrease coordination and communication costs. Innovation projects should involve members from all units from the beginning and include all units in all stages of an innovation project. Second, managers should strive to keep coordination and communication costs low especially when they face technological uncertainty. Technological complex environments might constantly raise the need for rebalancing exploration and exploitation activities and reorganization, thereby constantly increasing costs in all stages of the innovation process. As technological uncertainty negatively affects the diversity-innovation relationship in all stages, managers should pay special attention to it and install mechanisms in order to monitor technological developments closely and to act appropriately by tailoring the right technologies for the market. We want to emphasize that screening markets and technologies at the same time is important. If technological monitoring and market research are not located in the same departments (for example, if they are located in the IP and the marketing department instead), there is clearly the need to integrate these functions at least informally. Specialization would increase innovation success on the market, but probably at the cost of producing less patents that might be important to enrich the knowledge base of a firm and enable a firm to choose from a variety of technologies at the right time. Therefore, we do not consider it a valuable strategy.

Third, firms should think about a strategy in order to balance exploration and exploitation within their organization in a way that fits with their market and technological conditions and innovation process. If we suppose that there are units that are dealing predominantly with exploration and units that are dealing with exploitation, then regular contacts and tasks that are shared across units might be a good starting point in order to start implementing a process that aims at achieving a balance between the innovation activities and the corresponding units within a firm.

Further Research

Our paper highlights a number of interesting mechanisms how technological diversity affects different stages of the innovation process. However, additional work is needed to fully understand the complex relationships. In contrast to many other studies, our data comprises several industries and years. However, there is clearly the need for collecting data from several countries as the respective industry structure might influence the innovation processes and the kind of uncertainties the firms face. Further research should also try to get more nuanced data on inner-firm organization of innovation projects in order to show the whole picture of exploration and exploitation and how the different tasks are exactly assigned within an organization. Furthermore, research needs to tackle how technological uncertainty exactly arises and to show how firms can react adequately in order to select and develop the technologies that are most valuable for the market.

Appendix

3.A Tables

Table 3.A.1: Description of Variables

Variable	Description
Dependent Variables	
R&D Intensity	Research and Development (R&D) expenditures divided by sales in %
New Patent Applications	Number of patent applications in t*
Sales Share Innovative Products	Share of sales generated by new or improved products in %
Sales Share New Products	Share of sales generated by new products in %
Sales Share Improved Products	Share of sales generated by improved products in %
Main Explanatory Variables	
1-Herfindahl	One minus the Herfindahl index of patent sections in (t-1) and (t-2)*
Entropy	Total entropy across patent sections and classes in (t-1) and (t-2)*
Related Entropy	Entropy across patent sections in (t-1) and (t-2)*
Unrelated Entropy	Entropy within patent sections in (t-1) and (t-2)*
Uncertainty	Dummy variable that takes the value 1 if the relevance of technological uncertainty as an obstacle to innovation is high (4 or 5 on a 5 point Likert scale) and 0 otherwise
Control Variables	
Singlepatent	Dummy variable that takes the value 1 if the patent stock is based on a single patent and 0 otherwise
Patdummy	Dummy variable that takes the value 1 if the Patent stock is nonzero and 0 otherwise
Patent Stock	Number of patent applications in (t-1) and (t-2)*
Protection	Dummy variable that takes the value 1 if the effectiveness of innovation protection, e.g. through patents, copyrights, secrecy, is high (4 or 5 on a 5 point Likert scale) and 0 otherwise
Incoming Spillovers Customers	Dummy variable that takes the value 1 if the relevance of incoming spillovers from customers for innovation, is high (4 or 5 on a 5 point Likert scale) and 0 otherwise
Incoming Spillovers Suppliers	Dummy variable that takes the value 1 if the relevance of incoming spillovers from suppliers for innovation, is high (4 or 5 on a 5 point Likert scale) and 0 otherwise
Incoming Spillovers Competitors	Dummy variable that takes the value 1 if the relevance of incoming spillovers from competitors for innovation, is high (4 or 5 on a 5 point Likert scale) and 0 otherwise
R&D Cooperation	Dummy variable that takes the value 1 if the firm has R&D cooperation activities and 0 otherwise
Non-Price Comp	Intensity of non-price competition on a 5 point Likert scale
Price Comp	Intensity of price competition on a 5 point Likert scale
Qualification	Share of personnel with tertiary education in %
Size	Number of employees (full-time equivalents) in 1000
Technological Potential	Technological Potential on a 5 point Likert scale

* t refers to a three-year period

Table 3.A.2: Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Dependent Variables					
R&D Intensity	3,057	1.94	4.36	0	55.56
New Patent Applications	3,057	2.83	23.09	0	692
Sales Share Innovative Products	3,057	25.44	28.67	0	100
Sales Share New Products	3,057	11.92	17.84	0	100
Sales Share Improved Products	3,057	13.51	18.69	0	100
Main Explanatory Variables					
1-Herfindahl	3,057	0.12	0.25	0	0.93
Entropy	3,057	0.23	0.52	0	3.02
Related Entropy	3,057	0.09	0.26	0	1.8
Unrelated Entropy	3,057	0.14	0.33	0	1.65
Uncertainty	2,904	0.24	0.43	0	1
Control Variables					
Patdummy	3,057	0.28	0.45	0	1
Singlepatent	3,057	0.08	0.27	0	1
Patent Stock	3,057	3.47	33.02	0	1446.83
Protection	3,057	0.21	0.4	0	1
Incoming Spillovers Customers	3,057	0.5	0.5	0	1
Incoming Spillovers Suppliers	3,057	0.21	0.41	0	1
Incoming Spillovers Competitors	3,057	0.31	0.46	0	1
R&D Cooperation	3,057	0.27	0.44	0	1
Non-Price Comp	3,057	3.21	0.97	1	5
Price Comp	3,057	3.99	1.01	1	5
Size	3,057	0.21	0.76	0.001	21
Qualification	3,057	4.52	7.98	0	90
Technological Potential	3,057	2.96	1.08	1	5

Table 3.A.3: Main Results

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Innovative Products (%)	
Estimation method	OLS	FE	Poisson	FE Poisson	OLS	FE
Diversity Measure						
1-Herfindahl	0.896*	-0.815	1.515	0.438	-5.555	-10.518*
	(0.532)	(0.914)	(1.061)	(0.335)	(3.871)	(5.795)
Entropy	0.401	-0.246	1.916***	0.324**	-2.122	-5.271*
	(0.260)	(0.501)	(0.561)	(0.156)	(1.691)	(2.706)
Related Entropy	0.083	-0.265	2.270***	-0.095	-3.462	-11.168***
	(0.409)	(0.731)	(0.691)	(0.212)	(2.806)	(3.918)
Unrelated Entropy	0.650	-0.231	1.539**	0.706***	-1.076	-0.866
	(0.426)	(0.641)	(0.756)	(0.213)	(2.281)	(3.508)
N	3057	1429	3057	763	3057	1534

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level in the linear estimation and robust standard errors in the fixed effect estimation in parentheses. N refers to the number of observations, except for the FE poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively. All estimations include control variables as defined in Table 3.A.1 and time dummies.

Table 3.A.4: Technological Uncertainty

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Innovative Products (%)	
Estimation	OLS	FE	Poisson	FE Poisson	OLS	FE
Diversity Measure						
1-Herfindahl	0.655	-0.470	1.248	0.670*	-4.989	-9.745
	(0.606)	(0.939)	(1.370)	(0.352)	(4.132)	(6.009)
Uncertainty	0.237	0.548**	0.599	0.470**	-0.561	3.956**
	(0.221)	(0.263)	(0.726)	(0.194)	(1.348)	(1.822)
Interaction	0.344	-1.037	0.392	-0.655*	-1.692	-9.837*
	(0.763)	(0.986)	(1.447)	(0.358)	(4.243)	(5.412)
Entropy	0.332	-0.072	1.868***	0.420***	-1.930	-4.853*
	(0.308)	(0.551)	(0.688)	(0.162)	(1.848)	(2.767)
Uncertainty	0.271	0.546**	0.386	0.382**	-0.680	3.872**
	(0.219)	(0.260)	(0.689)	(0.162)	(1.322)	(1.782)
Interaction	0.062	-0.527	0.111	-0.213*	-0.509	-4.753*
	(0.366)	(0.458)	(0.559)	(0.109)	(1.988)	(2.596)
Related Entropy	-0.191	-0.053	2.224***	-0.159	-4.492	-12.271***
	(0.466)	(0.793)	(0.875)	(0.218)	(3.414)	(4.459)
Unrelated Entropy	0.723	-0.088	1.517	0.987***	-0.030	0.390
	(0.496)	(0.700)	(1.057)	(0.216)	(2.652)	(3.622)
Uncertainty	0.275	0.547**	0.361	0.477***	-0.652	3.935**
	(0.219)	(0.262)	(0.691)	(0.165)	(1.325)	(1.800)
Interaction Related	0.386	-0.478	-0.012	0.047	2.034	-1.195
	(0.858)	(0.869)	(0.847)	(0.144)	(4.404)	(4.491)
Interaction Unrelated	-0.162	-0.564	0.213	-0.539***	-2.327	-7.482*
	(0.741)	(0.834)	(0.859)	(0.202)	(3.789)	(4.462)
N	2904	1394	2904	746	2904	1490

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Interaction refers to the interaction of uncertainty and technological diversity. Estimates for R&D Intensity and Sales Share with Innovative Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level in the linear estimation and robust standard errors in the fixed effect estimation in parentheses. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively. All estimations include control variables as defined in Table 3.A.1 and time dummies.

3.B Further Tables

Table 3.B.1: Cross-Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 R&D Intensity	1.00																					
2 Patents	0.13	1.00																				
3 Inno Sales	0.37	0.08	1.00																			
4 New Sales	0.33	0.07	0.77	1.00																		
5 Improved Sales	0.25	0.05	0.80	0.23	1.00																	
6 Diversity	0.25	0.26	0.20	0.15	0.17	1.00																
7 Entropy	0.24	0.34	0.20	0.15	0.16	0.98	1.00															
8 Related Entropy	0.20	0.34	0.16	0.13	0.13	0.82	0.86	1.00														
9 Unrelated Entropy	0.23	0.28	0.19	0.14	0.16	0.91	0.92	0.59	1.00													
10 Technological Uncertainty	0.09	0.06	0.06	0.05	0.05	0.13	0.13	0.12	0.12	1.00												
11 Patdummy	0.25	0.19	0.25	0.18	0.21	0.77	0.72	0.59	0.68	0.14	1.00											
12 Singlepatent	0.06	-0.03	0.09	0.05	0.10	0.06	0.01	0.01	0.01	0.07	0.47	1.00										
13 Patent Stock	0.14	0.89	0.06	0.07	0.04	0.23	0.30	0.30	0.25	0.05	0.17	-0.02	1.00									
14 Protection	0.10	0.11	0.07	0.07	0.04	0.13	0.14	0.11	0.13	0.12	0.14	0.03	0.11	1.00								
15 Incoming Spillovers Customers	0.05	0.03	0.12	0.10	0.08	0.12	0.11	0.09	0.11	0.07	0.11	0.02	0.02	0.11	1.00							
16 Incoming Spillovers Suppliers	-0.07	-0.02	-0.02	0.00	-0.03	-0.06	-0.06	-0.05	-0.05	0.01	-0.07	-0.01	-0.03	0.07	0.10	1.00						
17 Incoming Spillovers Competitors	0.02	0.04	-0.01	0.00	-0.02	0.04	0.04	0.03	0.03	0.06	0.04	-0.01	0.02	0.08	0.21	0.10	1.00					
18 R&D Cooperation	0.29	0.12	0.34	0.27	0.26	0.26	0.26	0.22	0.24	0.09	0.29	0.11	0.10	0.08	0.09	0.02	0.02	1.00				
19 Non-Price Comp	0.11	0.04	0.14	0.13	0.09	0.08	0.08	0.07	0.07	0.05	0.09	0.05	0.04	0.09	0.11	0.07	0.08	0.11	1.00			
20 Price Comp	-0.02	0.01	0.04	0.02	0.05	0.06	0.06	0.04	0.07	0.02	0.08	0.05	0.01	0.01	0.07	0.09	0.09	0.07	0.06	1.00		
21 Size	0.11	0.56	0.12	0.10	0.09	0.29	0.35	0.34	0.29	0.06	0.20	-0.02	0.52	0.15	0.03	-0.01	0.07	0.17	0.06	0.03	1.00	
22 Qualification	0.38	0.11	0.18	0.14	0.15	0.17	0.17	0.14	0.17	0.06	0.16	0.05	0.10	0.09	0.02	-0.08	0.00	0.21	0.07	-0.06	0.17	1.00
23 Technological Potential	0.23	0.09	0.25	0.20	0.19	0.21	0.20	0.17	0.19	0.14	0.22	0.06	0.08	0.21	0.18	0.12	0.13	0.25	0.16	0.07	0.15	0.23

Table 3.B.2: Full Estimates of Main Estimates for Diversity - 1-Herfindahl

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Inno		Sales Share New		Sales Share Improved	
Estimations	OLS	FE	Poisson	FE Poisson	OLS	FE	OLS	FE	OLS	FE
1-Herfindahl	0.896*	-0.815	1.515	0.438	-5.555	-10.518*	-4.079	-5.854	-1.476	-4.665
	(0.532)	(0.914)	(1.061)	(0.335)	(3.871)	(5.795)	(2.559)	(4.439)	(2.682)	(4.995)
Singlepatent	-0.643*	-0.532	-5.394***	-0.452*	-1.941	0.896	-2.510	-1.133	0.569	2.029
	(0.379)	(0.552)	(0.913)	(0.246)	(2.448)	(3.543)	(1.607)	(2.780)	(1.811)	(3.351)
Patdummy	1.070***	0.664	10.000***	0.216	8.782***	1.624	4.226**	0.187	4.555***	1.437
	(0.319)	(0.506)	(1.369)	(0.287)	(2.470)	(3.916)	(1.655)	(2.749)	(1.651)	(3.750)
Incoming Spillovers Customers	-0.070	0.028	0.762	0.082	3.635***	1.604	2.020***	-0.216	1.616**	1.821*
	(0.148)	(0.195)	(0.734)	(0.112)	(0.997)	(1.431)	(0.643)	(1.006)	(0.684)	(1.066)
Incoming Spillovers Suppliers	-0.527***	0.155	-0.043	-0.074	-1.237	0.723	-0.074	0.752	-1.164	-0.028
	(0.157)	(0.167)	(0.627)	(0.126)	(1.153)	(1.689)	(0.772)	(1.211)	(0.754)	(1.194)
Incoming Spillovers Competitors	0.023	-0.014	0.083	-0.098	-3.681***	-3.910***	-1.486**	-0.914	-2.195***	-2.996**
	(0.145)	(0.186)	(0.625)	(0.087)	(1.012)	(1.491)	(0.649)	(0.966)	(0.710)	(1.194)
RnD Cooperation	1.691***	0.560*	0.137	0.024	11.759***	3.918**	5.606***	2.433*	6.152***	1.485
	(0.227)	(0.312)	(0.444)	(0.111)	(1.320)	(1.922)	(0.888)	(1.279)	(0.892)	(1.447)
Patent Stock	0.010**	0.021	0.010***	0.001	-0.032***	0.010	-0.009	0.013	-0.023***	-0.002
	(0.005)	(0.016)	(0.003)	(0.001)	(0.011)	(0.101)	(0.011)	(0.069)	(0.007)	(0.054)
Non-Price Comp	0.179**	-0.024	0.022	0.106**	1.557***	0.105	1.133***	0.569	0.424	-0.464
	(0.090)	(0.106)	(0.376)	(0.052)	(0.507)	(0.717)	(0.327)	(0.515)	(0.340)	(0.511)
Price Comp	-0.133*	-0.083	-0.265	0.089	0.433	-0.057	-0.100	0.299	0.533*	-0.356
	(0.077)	(0.110)	(0.287)	(0.069)	(0.471)	(0.756)	(0.309)	(0.467)	(0.320)	(0.628)
Size	-0.747**	0.074	2.051***	0.225	2.925*	10.683*	1.422	6.311*	1.503	4.373
	(0.376)	(0.866)	(0.285)	(0.164)	(1.557)	(6.118)	(1.213)	(3.693)	(1.165)	(4.632)
Size2	0.035	-0.019	-0.139***	-0.015	-0.106	-0.328	-0.046	-0.204	-0.060	-0.125
	(0.023)	(0.037)	(0.037)	(0.010)	(0.081)	(0.264)	(0.063)	(0.159)	(0.055)	(0.200)
Qualification	0.158***	0.071	0.021	-0.006	0.029	0.014	-0.042	-0.002	0.071	0.015
	(0.022)	(0.053)	(0.021)	(0.005)	(0.100)	(0.115)	(0.063)	(0.082)	(0.065)	(0.102)
Protection	0.201	0.324	0.491	-0.080	-1.596	-1.476	-0.418	0.561	-1.177	-2.037
	(0.194)	(0.206)	(0.503)	(0.107)	(1.200)	(1.646)	(0.788)	(1.163)	(0.878)	(1.268)
Technological Potential	0.374***	-0.018	-0.143	0.009	2.735***	1.244*	1.322***	0.371	1.413***	0.873*
	(0.071)	(0.090)	(0.301)	(0.070)	(0.503)	(0.673)	(0.342)	(0.475)	(0.341)	(0.520)
RnD Intensity			0.044	0.015	1.709***	0.506**	1.062***	0.375*	0.648***	0.132
			(0.056)	(0.011)	(0.162)	(0.204)	(0.132)	(0.210)	(0.125)	(0.226)
N	3057	3057	3057	763	3057	3057	3057	3057	3057	3057
N groups		1873		279		1873		1873		1873
F	19.99***	36.57***			37.02***	9.42***	19.07***	4.15***	21.07***	11.72***
Wald chi2			5745.16***	73.65***						

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products, Sales Share with New Products and Sales Share with Improved Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level and robust standard errors in the fixed effect estimation in parentheses. All estimations contain time dummies. Constants are suppressed. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 3.B.3: Full Estimates of Main Estimates for Diversity - Entropy

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Inno		Sales Share New		Sales Share Improved	
Estimations	OLS	FE	Poisson	FE Poisson	OLS	FE	OLS	FE	OLS	FE
Entropy	0.401	-0.246	1.916***	0.324**	-2.122	-5.271*	-1.692	-3.295*	-0.430	-1.976
	(0.260)	(0.501)	(0.561)	(0.156)	(1.691)	(2.706)	(1.144)	(1.962)	(1.161)	(2.334)
Singlepatent	-0.653*	-0.436	-4.339***	-0.373	-1.611	0.999	-2.365	-1.273	0.753	2.272
	(0.385)	(0.594)	(0.780)	(0.246)	(2.412)	(3.460)	(1.603)	(2.656)	(1.787)	(3.213)
Patdummy	1.129***	0.484	8.615***	0.103	8.065***	1.077	3.828**	0.188	4.237***	0.889
	(0.309)	(0.539)	(1.097)	(0.275)	(2.243)	(3.569)	(1.526)	(2.477)	(1.499)	(3.436)
Incoming Spillovers Customers	-0.071	0.025	0.476	0.078	3.630***	1.633	2.020***	-0.190	1.611**	1.823*
	(0.148)	(0.196)	(0.614)	(0.113)	(0.997)	(1.430)	(0.643)	(1.007)	(0.684)	(1.070)
Incoming Spillovers Suppliers	-0.529***	0.159	-0.454	-0.063	-1.221	0.738	-0.062	0.754	-1.159	-0.017
	(0.157)	(0.168)	(0.510)	(0.128)	(1.154)	(1.686)	(0.772)	(1.211)	(0.754)	(1.194)
Incoming Spillovers Competitors	0.025	-0.010	0.290	-0.053	-3.683***	-3.980***	-1.491**	-0.973	-2.192***	-3.007**
	(0.145)	(0.183)	(0.596)	(0.082)	(1.013)	(1.494)	(0.649)	(0.963)	(0.710)	(1.197)
RnD Cooperation	1.692***	0.566*	0.072	0.020	11.755***	3.926**	5.605***	2.427*	6.150***	1.499
	(0.227)	(0.312)	(0.426)	(0.108)	(1.321)	(1.919)	(0.888)	(1.274)	(0.892)	(1.448)
Patent Stock	0.010*	0.022	0.009***	0.001	-0.030**	0.029	-0.007	0.025	-0.022***	0.005
	(0.005)	(0.016)	(0.002)	(0.001)	(0.012)	(0.097)	(0.011)	(0.067)	(0.007)	(0.051)
Non-Price Comp	0.178**	-0.021	0.045	0.088*	1.561***	0.141	1.137***	0.589	0.424	-0.449
	(0.090)	(0.106)	(0.296)	(0.051)	(0.507)	(0.715)	(0.327)	(0.514)	(0.340)	(0.511)
Price Comp	-0.133*	-0.083	-0.243	0.084	0.432	-0.055	-0.099	0.301	0.532*	-0.355
	(0.077)	(0.110)	(0.271)	(0.067)	(0.471)	(0.755)	(0.309)	(0.466)	(0.320)	(0.628)
Size	-0.776**	0.061	1.608***	0.228	3.032*	10.777*	1.528	6.406*	1.504	4.370
	(0.381)	(0.881)	(0.253)	(0.160)	(1.569)	(6.119)	(1.234)	(3.669)	(1.189)	(4.670)
Size2	0.036	-0.018	-0.114***	-0.015	-0.113	-0.332	-0.053	-0.208	-0.060	-0.124
	(0.023)	(0.038)	(0.031)	(0.010)	(0.082)	(0.264)	(0.064)	(0.158)	(0.056)	(0.202)
Qualification	0.158***	0.071	0.018	-0.006	0.029	0.017	-0.042	0.001	0.071	0.016
	(0.022)	(0.053)	(0.019)	(0.005)	(0.100)	(0.115)	(0.063)	(0.082)	(0.065)	(0.102)
Protection	0.198	0.328	0.527	-0.073	-1.574	-1.478	-0.404	0.551	-1.170	-2.029
	(0.195)	(0.204)	(0.478)	(0.103)	(1.200)	(1.645)	(0.788)	(1.160)	(0.878)	(1.268)
Technological Potential	0.375***	-0.020	-0.125	0.003	2.727***	1.229*	1.316***	0.363	1.411***	0.866*
	(0.071)	(0.090)	(0.305)	(0.066)	(0.502)	(0.673)	(0.342)	(0.474)	(0.341)	(0.519)
RnD Intensity			0.050	0.013	1.708***	0.510**	1.061***	0.376*	0.647***	0.134
			(0.046)	(0.010)	(0.162)	(0.204)	(0.132)	(0.209)	(0.124)	(0.226)
N	3057	3057	3057	763	3057	3057	3057	3057	3057	3057
N groups		1873		279		1873		1873		1873
F	19.93***	35.78***			37.00***	9.53***	18.90***	4.23***	21.07***	11.86***
Wald chi2			6382.64***	75.98***						

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products, Sales Share with New Products and Sales Share with Improved Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level and robust standard errors in the fixed effect estimation in parentheses. All estimations contain time dummies. Constants are suppressed. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 3.B.4: Full Estimates of Main Estimates for Diversity - Related and Unrelated Entropy

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Inno		Sales Share New		Sales Share Improved	
Estimations	OLS	FE	Poisson	FE Poisson	OLS	FE	OLS	FE	OLS	FE
Entropy Related	0.083	-0.265	2.270***	-0.095	-3.462	-11.168***	-1.282	-4.946*	-2.180	-6.222*
	(0.409)	(0.731)	(0.691)	(0.212)	(2.806)	(3.918)	(1.739)	(2.774)	(1.831)	(3.355)
Entropy Unrelated	0.650	-0.231	1.539**	0.706***	-1.076	-0.866	-2.012	-2.062	0.937	1.195
	(0.426)	(0.641)	(0.756)	(0.213)	(2.281)	(3.508)	(1.568)	(2.663)	(1.604)	(2.934)
Singlepatent	-0.631	-0.434	-4.400***	-0.342	-1.519	1.616	-2.393	-1.100	0.873	2.715
	(0.387)	(0.600)	(0.793)	(0.258)	(2.418)	(3.501)	(1.606)	(2.698)	(1.795)	(3.266)
Patdummy	1.103***	0.482	8.692***	0.003	7.958***	0.525	3.861**	0.033	4.097***	0.492
	(0.315)	(0.546)	(1.108)	(0.266)	(2.239)	(3.575)	(1.531)	(2.496)	(1.501)	(3.454)
Incoming Spillovers Customers	-0.073	0.025	0.480	0.091	3.621***	1.599	2.022***	-0.199	1.599**	1.798*
	(0.149)	(0.196)	(0.609)	(0.109)	(0.998)	(1.433)	(0.643)	(1.008)	(0.684)	(1.069)
Incoming Spillovers Suppliers	-0.530***	0.159	-0.466	-0.066	-1.226	0.592	-0.060	0.714	-1.165	-0.122
	(0.157)	(0.167)	(0.503)	(0.122)	(1.154)	(1.677)	(0.772)	(1.212)	(0.754)	(1.185)
Incoming Spillovers Competitors	0.026	-0.010	0.295	-0.070	-3.680***	-3.949***	-1.492**	-0.964	-2.188***	-2.985**
	(0.145)	(0.183)	(0.594)	(0.086)	(1.013)	(1.492)	(0.650)	(0.962)	(0.710)	(1.199)
RnD Cooperation	1.693***	0.566*	0.073	0.037	11.762***	3.844**	5.602***	2.404*	6.160***	1.440
	(0.227)	(0.313)	(0.425)	(0.108)	(1.321)	(1.921)	(0.888)	(1.273)	(0.893)	(1.447)
Patent Stock	0.010**	0.022	0.009***	0.001	-0.029**	0.038	-0.008	0.027	-0.021***	0.011
	(0.005)	(0.016)	(0.002)	(0.001)	(0.012)	(0.096)	(0.011)	(0.066)	(0.006)	(0.051)
Non-Price Comp	0.179**	-0.021	0.051	0.106**	1.564***	0.165	1.136***	0.596	0.428	-0.431
	(0.090)	(0.106)	(0.290)	(0.052)	(0.507)	(0.717)	(0.327)	(0.513)	(0.340)	(0.513)
Price Comp	-0.135*	-0.083	-0.247	0.074	0.426	-0.094	-0.097	0.290	0.523	-0.384
	(0.077)	(0.110)	(0.270)	(0.067)	(0.472)	(0.757)	(0.310)	(0.468)	(0.320)	(0.628)
Size	-0.762**	0.060	1.587***	0.225	3.091**	10.323*	1.510	6.279*	1.581	4.044
	(0.379)	(0.886)	(0.249)	(0.141)	(1.574)	(6.007)	(1.233)	(3.641)	(1.198)	(4.656)
Size2	0.036	-0.018	-0.112***	-0.014*	-0.116	-0.310	-0.052	-0.202	-0.064	-0.108
	(0.023)	(0.038)	(0.030)	(0.008)	(0.082)	(0.259)	(0.064)	(0.157)	(0.057)	(0.201)
Qualification	0.158***	0.071	0.020	-0.009*	0.028	0.007	-0.042	-0.002	0.070	0.009
	(0.022)	(0.053)	(0.018)	(0.005)	(0.100)	(0.116)	(0.063)	(0.083)	(0.065)	(0.102)
Protection	0.193	0.328	0.524	-0.093	-1.592	-1.638	-0.398	0.506	-1.194	-2.145*
	(0.194)	(0.207)	(0.478)	(0.104)	(1.202)	(1.635)	(0.787)	(1.143)	(0.878)	(1.257)
Technological Potential	0.375***	-0.020	-0.101	-0.009	2.726***	1.222*	1.316***	0.361	1.410***	0.861*
	(0.071)	(0.091)	(0.311)	(0.065)	(0.502)	(0.671)	(0.342)	(0.475)	(0.341)	(0.518)
RnD Intensity			0.051	0.015	1.707***	0.509**	1.062***	0.376*	0.645***	0.133
			(0.046)	(0.010)	(0.162)	(0.204)	(0.131)	(0.209)	(0.124)	(0.226)
N	3057	3057	3057	763	3057	3057	3057	3057	3057	3057
N groups		1873		279		1873		1873		1873
F	19.20***	33.90***								
Wald chi2			6588.62***	92.19***	35.20***	8.58***	18.11***	4.03***	20.05***	10.59***

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products, Sales Share with New Products and Sales Share with Improved Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level and robust standard errors in the fixed effect estimation in parentheses. All estimations contain time dummies. Constants are suppressed. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 3.B.5: Full Estimates of Main Estimates for Diversity & Uncertainty - 1-Herfindahl

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Inno		Sales Share New		Sales Share Improved	
Estimations	OLS	FE	Poisson	FE Poisson	OLS	FE	OLS	FE	OLS	FE
1-Herfindahl	0.655	-0.470	1.248	0.670*	-4.989	-9.745	-3.663	-7.870*	-1.326	-1.875
	(0.606)	(0.939)	(1.370)	(0.352)	(4.132)	(6.009)	(2.782)	(4.449)	(2.913)	(5.111)
Uncertainty	0.237	0.548**	0.599	0.470**	-0.561	3.956**	-0.158	0.973	-0.403	2.983**
	(0.221)	(0.263)	(0.726)	(0.194)	(1.348)	(1.822)	(0.837)	(1.084)	(0.942)	(1.480)
Interaction	0.344	-1.037	0.392	-0.655*	-1.692	-9.837*	-1.773	-1.380	0.081	-8.458*
	(0.763)	(0.986)	(1.447)	(0.358)	(4.243)	(5.412)	(2.780)	(3.666)	(2.991)	(4.570)
Singlepatent	-0.724*	-0.615	-5.594***	-0.342	-2.005	1.120	-2.693*	-1.426	0.687	2.546
	(0.377)	(0.563)	(0.950)	(0.233)	(2.461)	(3.564)	(1.597)	(2.695)	(1.836)	(3.364)
Patdummy	1.095***	0.694	10.286***	0.145	8.660***	1.988	4.200**	0.763	4.460***	1.224
	(0.317)	(0.507)	(1.433)	(0.289)	(2.481)	(3.863)	(1.669)	(2.660)	(1.669)	(3.810)
Incoming Spillovers Customers	-0.110	0.010	0.829	0.081	3.784***	1.194	2.122***	-0.526	1.662**	1.720
	(0.152)	(0.207)	(0.743)	(0.107)	(1.030)	(1.487)	(0.668)	(1.054)	(0.709)	(1.112)
Incoming Spillovers Suppliers	-0.539***	0.187	-0.159	-0.064	-1.333	1.681	-0.086	1.400	-1.247	0.281
	(0.164)	(0.174)	(0.600)	(0.121)	(1.184)	(1.737)	(0.784)	(1.220)	(0.786)	(1.254)
Incoming Spillovers Competitors	0.052	-0.027	0.137	-0.121	-3.475***	-3.838**	-1.367**	-1.092	-2.108***	-2.746**
	(0.149)	(0.200)	(0.616)	(0.088)	(1.043)	(1.547)	(0.672)	(1.018)	(0.733)	(1.256)
RnD Cooperation	1.602***	0.567*	0.165	0.044	11.017***	3.665*	5.276***	2.357*	5.740***	1.307
	(0.225)	(0.318)	(0.470)	(0.110)	(1.326)	(1.938)	(0.885)	(1.298)	(0.903)	(1.453)
Patent Stock	0.010**	0.021	0.010***	0.002	-0.031***	0.021	-0.009	0.014	-0.022***	0.007
	(0.005)	(0.016)	(0.003)	(0.001)	(0.012)	(0.097)	(0.011)	(0.069)	(0.007)	(0.051)
Non-Price Comp	0.177*	-0.031	-0.009	0.132**	1.496***	0.168	1.101***	0.639	0.395	-0.471
	(0.096)	(0.112)	(0.382)	(0.054)	(0.530)	(0.748)	(0.341)	(0.502)	(0.358)	(0.536)
Price Comp	-0.140*	-0.100	-0.311	0.091	0.407	0.002	-0.104	0.561	0.511	-0.559
	(0.082)	(0.118)	(0.304)	(0.069)	(0.501)	(0.784)	(0.331)	(0.482)	(0.340)	(0.634)
Size	-0.738*	0.093	2.228***	0.253	2.947*	12.085*	1.448	7.197*	1.499	4.888
	(0.382)	(0.891)	(0.300)	(0.157)	(1.569)	(6.445)	(1.222)	(3.746)	(1.175)	(4.919)
Size2	0.033	-0.020	-0.153***	-0.016	-0.106	-0.385	-0.046	-0.241	-0.060	-0.143
	(0.023)	(0.038)	(0.037)	(0.010)	(0.081)	(0.278)	(0.063)	(0.161)	(0.056)	(0.212)
Qualification	0.160***	0.073	0.021	-0.005	0.032	0.024	-0.036	-0.004	0.068	0.028
	(0.022)	(0.054)	(0.022)	(0.005)	(0.102)	(0.116)	(0.065)	(0.083)	(0.067)	(0.103)
Protection	0.201	0.311	0.524	-0.093	-1.346	-1.958	-0.295	0.637	-1.051	-2.595**
	(0.210)	(0.221)	(0.521)	(0.106)	(1.259)	(1.702)	(0.824)	(1.153)	(0.928)	(1.277)
Technological Potential	0.384***	-0.024	-0.129	0.006	2.859***	1.145*	1.315***	0.304	1.543***	0.840
	(0.076)	(0.098)	(0.322)	(0.067)	(0.521)	(0.695)	(0.357)	(0.482)	(0.355)	(0.546)
RnD Intensity			0.048	0.014	1.669***	0.443**	1.054***	0.360*	0.615***	0.084
			(0.057)	(0.011)	(0.161)	(0.202)	(0.135)	(0.210)	(0.125)	(0.229)
N	2904	2904	2904	746	2904	2904	2904	2904	2904	2904
N groups		1795		272		1795		1795		1795
F	17.82***	23.96***			31.25***	8.94***	15.97***	4.45***	17.74***	9.39***
Wald chi2			7009.07***	96.05***						

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products, Sales Share with New Products and Sales Share with Improved Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level and robust standard errors in the fixed effect estimation in parentheses. All estimations contain time dummies. Constants are suppressed. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 3.B.6: Full Estimates of Main Estimates for Diversity & Uncertainty - Entropy

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Inno		Sales Share New		Sales Share Improved	
Estimations	OLS	FE	Poisson	FE Poisson	OLS	FE	OLS	FE	OLS	FE
Entropy	0.332	-0.072	1.868***	0.420***	-1.930	-4.853*	-1.544	-4.173**	-0.386	-0.679
	(0.308)	(0.551)	(0.688)	(0.162)	(1.848)	(2.767)	(1.279)	(2.003)	(1.302)	(2.409)
Uncertainty	0.271	0.546**	0.386	0.382**	-0.680	3.872**	-0.254	0.855	-0.426	3.017**
	(0.219)	(0.260)	(0.689)	(0.162)	(1.322)	(1.782)	(0.823)	(1.075)	(0.927)	(1.452)
Interaction	0.062	-0.527	0.111	-0.213*	-0.509	-4.753*	-0.598	-0.356	0.089	-4.397**
	(0.366)	(0.458)	(0.559)	(0.109)	(1.988)	(2.596)	(1.340)	(1.702)	(1.375)	(2.184)
Singlepatent	-0.727*	-0.520	-4.523***	-0.242	-1.666	1.305	-2.531	-1.411	0.865	2.716
	(0.386)	(0.607)	(0.830)	(0.235)	(2.430)	(3.464)	(1.602)	(2.605)	(1.814)	(3.234)
Patdummy	1.135***	0.516	8.887***	0.018	7.932***	1.230	3.772**	0.443	4.160***	0.787
	(0.312)	(0.543)	(1.172)	(0.278)	(2.259)	(3.503)	(1.545)	(2.433)	(1.519)	(3.493)
Incoming Spillovers Customers	-0.110	0.005	0.540	0.071	3.777***	1.210	2.119***	-0.494	1.657**	1.705
	(0.152)	(0.207)	(0.649)	(0.106)	(1.029)	(1.486)	(0.669)	(1.056)	(0.709)	(1.116)
Incoming Spillovers Suppliers	-0.541***	0.192	-0.536	-0.053	-1.318	1.713	-0.074	1.401	-1.244	0.312
	(0.164)	(0.175)	(0.519)	(0.120)	(1.185)	(1.735)	(0.784)	(1.221)	(0.786)	(1.254)
Incoming Spillovers Competitors	0.055	-0.026	0.339	-0.074	-3.481***	-3.944**	-1.376**	-1.159	-2.104***	-2.785**
	(0.149)	(0.196)	(0.594)	(0.081)	(1.043)	(1.550)	(0.672)	(1.015)	(0.733)	(1.260)
RnD Cooperation	1.601***	0.568*	0.089	0.045	11.017***	3.647*	5.278***	2.367*	5.739***	1.280
	(0.225)	(0.317)	(0.450)	(0.106)	(1.326)	(1.937)	(0.885)	(1.294)	(0.903)	(1.453)
Patent Stock	0.010**	0.022	0.009***	0.001	-0.029**	0.047	-0.007	0.029	-0.022***	0.018
	(0.005)	(0.016)	(0.002)	(0.001)	(0.012)	(0.091)	(0.011)	(0.066)	(0.007)	(0.047)
Non-Price Comp	0.176*	-0.027	0.033	0.111**	1.499***	0.226	1.104***	0.667	0.395	-0.441
	(0.096)	(0.112)	(0.306)	(0.051)	(0.530)	(0.746)	(0.342)	(0.501)	(0.358)	(0.535)
Price Comp	-0.141*	-0.100	-0.266	0.089	0.406	0.004	-0.104	0.562	0.510	-0.559
	(0.082)	(0.118)	(0.290)	(0.068)	(0.501)	(0.781)	(0.331)	(0.481)	(0.340)	(0.633)
Size	-0.767**	0.092	1.747***	0.262*	3.056*	12.292*	1.562	7.276*	1.494	5.015
	(0.386)	(0.900)	(0.303)	(0.150)	(1.581)	(6.465)	(1.245)	(3.715)	(1.199)	(4.943)
Size2	0.035	-0.019	-0.125***	-0.017	-0.113	-0.393	-0.053	-0.245	-0.060	-0.148
	(0.023)	(0.039)	(0.034)	(0.010)	(0.082)	(0.279)	(0.064)	(0.160)	(0.057)	(0.213)
Qualification	0.160***	0.073	0.019	-0.005	0.031	0.028	-0.036	-0.002	0.068	0.030
	(0.022)	(0.054)	(0.020)	(0.005)	(0.102)	(0.116)	(0.065)	(0.083)	(0.067)	(0.103)
Protection	0.194	0.317	0.551	-0.074	-1.310	-1.932	-0.268	0.644	-1.042	-2.576**
	(0.210)	(0.218)	(0.498)	(0.105)	(1.258)	(1.697)	(0.823)	(1.147)	(0.927)	(1.277)
Technological Potential	0.385***	-0.025	-0.116	0.007	2.853***	1.133	1.311***	0.296	1.542***	0.836
	(0.076)	(0.098)	(0.318)	(0.063)	(0.521)	(0.694)	(0.357)	(0.481)	(0.355)	(0.544)
RnD Intensity			0.054	0.012	1.668***	0.446**	1.053***	0.363*	0.614***	0.083
			(0.049)	(0.010)	(0.161)	(0.200)	(0.135)	(0.209)	(0.125)	(0.228)
N	2904	2904	2904	746	2904	2904	2904	2904	2904	2904
N groups		1795		272		1795		1795		1795
F	17.92***	22.29***			31.18***	8.89***			17.78***	9.28***
Wald chi2			8231.14***	110.89***			15.82***	4.56***		

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products, Sales Share with New Products and Sales Share with Improved Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level and robust standard errors in the fixed effect estimation in parentheses. All estimations contain time dummies. Constants are suppressed. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

Table 3.B.7: Full Estimates of Main Estimates for Diversity & Uncertainty - Related and Unrelated Entropy

Dependent Variable	R&D Intensity (%)		New Patent Applications		Sales Share Inno		Sales Share New		Sales Share Improved	
Estimations	OLS	FE	Poisson	FE Poisson	OLS	FE	OLS	FE	OLS	FE
Entropy Related	-0.191	-0.053	2.224**	-0.159	-4.492	-12.271***	-1.248	-6.205**	-3.243	-6.066
	(0.466)	(0.793)	(0.875)	(0.218)	(3.414)	(4.459)	(2.143)	(2.998)	(2.172)	(3.783)
Entropy Unrelated	0.723	-0.088	1.517	0.987***	-0.030	0.390	-1.759	-2.799	1.729	3.189
	(0.496)	(0.700)	(1.057)	(0.216)	(2.652)	(3.622)	(1.778)	(2.696)	(1.902)	(3.063)
Uncertainty	0.275	0.547**	0.361	0.477***	-0.652	3.935**	-0.260	0.902	-0.393	3.033**
	(0.219)	(0.262)	(0.691)	(0.165)	(1.325)	(1.800)	(0.824)	(1.083)	(0.929)	(1.462)
Interaction Related	0.386	-0.478	-0.012	0.047	2.034	-1.195	-1.141	2.176	3.175	-3.372
	(0.858)	(0.869)	(0.847)	(0.144)	(4.404)	(4.491)	(2.736)	(3.553)	(3.008)	(3.547)
Interaction Unrelated	-0.162	-0.564	0.213	-0.539***	-2.327	-7.482*	-0.199	-2.305	-2.128	-5.176
	(0.741)	(0.834)	(0.859)	(0.202)	(3.789)	(4.462)	(2.448)	(2.820)	(2.697)	(3.485)
Singlepatent	-0.702*	-0.525	-4.571***	-0.172	-1.572	1.877	-2.536	-1.362	0.963	3.239
	(0.388)	(0.617)	(0.845)	(0.232)	(2.435)	(3.535)	(1.608)	(2.665)	(1.818)	(3.284)
Patdummy	1.103***	0.519	8.953***	-0.125	7.803***	0.704	3.781**	0.360	4.023***	0.344
	(0.317)	(0.552)	(1.189)	(0.266)	(2.256)	(3.518)	(1.550)	(2.470)	(1.519)	(3.508)
Incoming Spillovers Customers	-0.110	0.005	0.536	0.101	3.781***	1.202	2.117***	-0.485	1.663**	1.686
	(0.152)	(0.208)	(0.648)	(0.099)	(1.032)	(1.493)	(0.669)	(1.058)	(0.710)	(1.117)
Incoming Spillovers Suppliers	-0.540***	0.194	-0.545	-0.037	-1.314	1.643	-0.076	1.427	-1.237	0.215
	(0.164)	(0.174)	(0.515)	(0.111)	(1.186)	(1.737)	(0.785)	(1.226)	(0.787)	(1.247)
Incoming Spillovers Competitors	0.055	-0.026	0.340	-0.147*	-3.487***	-3.942**	-1.374**	-1.167	-2.112***	-2.775**
	(0.149)	(0.196)	(0.594)	(0.085)	(1.043)	(1.551)	(0.673)	(1.016)	(0.734)	(1.263)
RnD Cooperation	1.605***	0.569*	0.087	0.063	11.042***	3.586*	5.274***	2.375*	5.769***	1.210
	(0.225)	(0.319)	(0.445)	(0.107)	(1.329)	(1.938)	(0.886)	(1.296)	(0.905)	(1.452)
Patent Stock	0.010**	0.022	0.009***	0.002	-0.028**	0.055	-0.007	0.031	-0.021***	0.024
	(0.005)	(0.016)	(0.002)	(0.001)	(0.012)	(0.090)	(0.011)	(0.066)	(0.006)	(0.046)
Non-Price Comp	0.176*	-0.027	0.039	0.145***	1.500***	0.236	1.104***	0.661	0.395	-0.425
	(0.096)	(0.112)	(0.303)	(0.051)	(0.530)	(0.750)	(0.341)	(0.501)	(0.358)	(0.537)
Price Comp	-0.142*	-0.100	-0.267	0.078	0.400	-0.021	-0.103	0.569	0.503	-0.589
	(0.082)	(0.119)	(0.289)	(0.066)	(0.501)	(0.782)	(0.331)	(0.482)	(0.340)	(0.632)
Size	-0.748*	0.097	1.724***	0.262**	3.131**	12.052*	1.557	7.334**	1.573	4.718
	(0.384)	(0.906)	(0.302)	(0.125)	(1.587)	(6.379)	(1.243)	(3.692)	(1.203)	(4.940)
Size2	0.034	-0.020	-0.124***	-0.015**	-0.118	-0.380	-0.053	-0.247	-0.066	-0.133
	(0.023)	(0.039)	(0.034)	(0.007)	(0.083)	(0.275)	(0.064)	(0.159)	(0.058)	(0.213)
Qualification	0.160***	0.073	0.020	-0.006	0.030	0.015	-0.036	-0.006	0.067	0.021
	(0.022)	(0.054)	(0.020)	(0.005)	(0.102)	(0.116)	(0.065)	(0.083)	(0.067)	(0.103)
Protection	0.188	0.318	0.554	-0.132	-1.336	-2.163	-0.266	0.574	-1.070	-2.738**
	(0.210)	(0.219)	(0.509)	(0.108)	(1.260)	(1.691)	(0.822)	(1.133)	(0.927)	(1.266)
Technological Potential	0.384***	-0.025	-0.096	-0.001	2.849***	1.113	1.312***	0.286	1.537***	0.828
	(0.076)	(0.099)	(0.323)	(0.062)	(0.520)	(0.690)	(0.357)	(0.481)	(0.354)	(0.542)
RnD Intensity			0.055	0.012	1.665***	0.446**	1.054***	0.362*	0.611***	0.084
			(0.049)	(0.010)	(0.161)	(0.200)	(0.135)	(0.209)	(0.125)	(0.229)
N	2904	2904	2904	746	2904	2904	2904	2904	2904	2904
N groups		1795		272		1795		1795		1795
F	16.68***	20.67***			28.78***	7.87***	14.65***	4.27***	16.60***	7.98***
Wald chi2			8425.30***	135.60***						

Each table block displays two individual estimations showing partial correlations and fixed effect estimates, respectively. Estimates for R&D Intensity and Sales Share with Innovative Products, Sales Share with New Products and Sales Share with Improved Products show OLS coefficient estimates and robust standard errors clustered at firm level in parentheses. Estimates for New Patent Applications display marginal effects of Poisson regressions with robust standard errors clustered at firm level and robust standard errors in the fixed effect estimation in parentheses. All estimations contain time dummies. Constants are suppressed. N refers to the number of observations, except for the FE Poisson estimates, where it refers to the number of observations for firms that have a change in the dependent variable. *,** and *** denote significance on the levels 10%, 5% and 1%, respectively.

4 What Determines International and Inter-sectoral Knowledge Flows? The Impact of Absorptive Capacity, Technological Distance and Spillovers *

4.1 Introduction

Knowledge integration across borders and industries is an important facet of globalization and digitalization where exchange of intangible goods and even tacit knowledge across geographic and institutional borders becomes more and more important. The process of innovation – that is to a large degree sequential – relies on firms learning from external knowledge (Cohen and Levinthal 1989) and 'recombining' new-to-the-firm knowledge in meaningful ways with own knowledge (Nelson and Winter 1982; Weitzman 1998). Combining knowledge across industries seems to be prevalent: For example, in the automobile industry, some 'disruptive trends' are based on business models and technologies that have their origin in other industries such as the ICT industry (McKinsey&Company 2016). Another example is the convergence of the pharmaceutical and food industry in order to produce new 'functional foods' (Hacklin et al. 2013).

External knowledge absorption partly depends on knowledge spillovers that are non-pecuniary externalities arising from others' knowledge activities and are empirically found to influence innovation and growth positively (see Coe and Helpman 1995; Griliches 1992; Jaffe 1986). The basic condition that spillovers can be absorbed is that a firm has enough absorptive capacity, i.e., that it is able to understand and exploit external knowledge and to apply it to commercial

*I use the first person plural although the paper is single-authored.

ends (Cohen and Levinthal 1989). Of course, knowledge that is absorbed from external sources might lead to further knowledge flows back to the initial sources or it might diffuse to other firms, industries or countries after having been enriched with own knowledge (if not kept secret), thereby contributing to the process of sequential innovation.

This paper tries to disentangle different sources and recipients of knowledge and to analyse their impact on these further knowledge flows. First, it analyses determinants of knowledge flows that might arise from prior knowledge exchange between firms, industries or countries or from knowledge spillovers. Second, it takes into account whether industries that engage in the process described above are high-tech or low-tech. According to convential wisdom, mainly high-tech industries are involved in inter-industry knowledge integration, at the same time high-tech sector-countries might act as learning sources for other less advanced sector-countries (see Griffith et al. 2004). Hence, the effects on further knowledge flows might depend to a large degree on whether the industries are high-tech or low-tech industries.

We study an input output framework using a gravity-like model with patent citations as proxies for knowledge flows and (weighted) R&D stocks as proxies for absorptive capacity and spillovers. In order to capture the direction and amount of knowledge flows between sector-countries, we emulate World Input Output Tables by counting the number of forward citations that patent applications of sector-countries (input dimension) receive from other sector-countries (output dimension) and control for industry- and country-specific factors. Beginning with Peri (2005) and Maurseth and Verspagen (2002), there are now more and more papers studying international knowledge flows in gravity models using patent citations as proxies for knowledge flows (e.g., Li 2014; Morescalchi et al. 2014).

The paper contributes to literature, both methodologically and conceptually: First, it looks at the process of knowledge generation and absorption and tries to disentangle multiple determinants of knowledge flows that might lead to more innovation and growth. The input-output framework makes it possible to distinguish between knowledge accumulated in the input and output sector-countries and spillovers from sector-countries that are external to the input and output sector-country. Second, it analyzes knowledge flows at a more detailed level than other papers on international knowledge flows, namely at the sector-country level with both an input and output dimension. The sectoral dimension is missing in existing literature to a large extent (see Badinger and Egger 2015) apart from a few notable exceptions (see Frantzen 2002; Keller 2002b; Malerba et al. 2013; Park 2004). Third, the paper also includes technological distance at the sector-country and sectoral level in addition to technological distance at country level and in contrast to literature on trade and knowledge flows that is interested in a geographical dimension only.

Our results indicate that in input-output relationships, knowledge accumulated in the output sector-country is the most important source for the generation of further knowledge flows, especially when compared with own knowledge accumulated by the input sector-country. We find that knowledge spillovers from external sector-countries outside the input-output relationship are also important, but whether this kind of knowledge is used to generate further knowledge flows depends on the technological advancement of the involved sector-countries. Fully external knowledge is mainly used by low-tech input sector-countries, but the preferred source for producing knowledge flows is the high-tech output sector-country's knowledge stock (i.e., the input sector-country is low-tech and the output sector-country high-tech). Technological distance between sector-countries is found to be a major impediment of future knowledge flows between sector-country pairs, but the degree of whether the distance between countries or the distance between industries matter again depends on the technological advancement of the respective sector-countries. In sum, the results show that knowledge flows depend to a large extent on prior knowledge exchange between sector-countries and not so much on external knowledge spillovers as suggested, but that the composition of the respective sector-country relationships greatly matters with respect to the relevance of knowledge absorption.[1]

The paper is structured as follows: Section 4.2 gives a literature review on absorptive capacity and knowledge spillovers and states our hypotheses. Section 4.3 presents our empirical model and the estimation strategy. Section 4.4 describes the data and variables. Section 4.5 discusses the results and Section 4.6 concludes.

4.2 Literature Review and Hypotheses

4.2.1 Absorptive Capacity and Spillovers

Absorptive capacity mainly depends on past experience with R&D activities and the stock of highly-educated engineers and inventors who are able to understand external knowledge and to apply it to commercial ends. In addition, there is a continuous inflow of knowledge that

[1] If we speak of knowledge exchange in this paper, we only refer to 'involuntary' knowledge exchange based on spillovers and not formal knowledge exchange that is based on cooperation or licensing agreements etc.

spills from competitors (intra-industry spillovers), other industries or countries. A country's or industry's ability to exploit this kind of knowledge requires enough innate absorptive capacity. In Cohen and Levinthal's framework, spillovers and absorptive capacity are positively related (the more spillovers, the more absorptive capacity is needed to acquire knowledge). In addition, R&D has two faces: It stimulates innovation and establishes absorptive capacity that enables an economic entity to utilize external knowledge. Hence, the higher the absorptive capacity of firms, industries or countries, the more knowledge spillover to them will take place (Hall et al. 2010). Once internalized, spillovers also add to compound absorptive capacity by broadening the knowledge base. In this way, they may enhance further knowledge flows and innovation.

There is still little known about the potential of absorptive capacity to increase future knowledge flows, i.e. the potential not only to attract knowledge from a spillover source but also to provide the base for further knowledge flows back to the initial spillover source or another external unit.[2] This potential impact of absorptive capacity refers to public benefits (generation of knowledge from which others can benefit) in contrast to the private commercial benefits that are typically studied. Indeed, an economic unit that is able to recognize and utilize relevant internal and external knowledge might not only generate own inventions (and corresponding profits) out of it, but might also provide the base for knowledge that again flows to firms, industries or countries.[3] Over time, the process of knowledge absorption combined with recombination, continuous improvement and enhancement of knowledge may trigger higher sequential innovative activity.

We distinguish between four potential sources of absorptive capacity and spillovers at sector-country level and examine their impact on the extent of future knowledge flows between a certain pair of sector-countries: First, the existing internal knowledge stock of an input sector-country, second, the knowledge stock of the output sector-country (that is the sector-country that cites the focal input sector-country, i.e. the sector-country that draws on the knowledge generated by the input sector-country later on), third, external knowledge spillovers from sector-countries that are external to the input sector-country (i.e., sector-countries that are not part of the input-output relationship, that do not cite the focal sector-country, but still can be used as source of spillovers initially), and, fourth, external knowledge spillovers from sector-countries

[2] One study that looks at knowledge flows as a function of absorptive capacity is Mukherji and Silberman (2013).

[3] However, in a sequential setting, knowledge flows back to the original inventor also give rise to private returns to innovation as an economic entity can benefit from the recombination of its past inventions with external ideas (Belenzon 2012; Yang et al. 2010).

that are external to the output sector-country.[4] Knowledge from the output sector-country might be especially relevant as this sector-country is the receiver of knowledge flows later on. Knowledge that is exchanged between input and output sector-countries might be tailored to the sector-countries' needs, whereas, in the case of fully external spillovers, the knowledge is rather unspecific which might shrink usability.

Conceptually, we distinguish between spillovers and knowledge flows in the following way: Spillovers are knowledge externalities proxied by the accumulated R&D expenditures of external sector-countries that occur involuntarily, whereas knowledge flows refer to voluntary but informal forms of knowledge exchange between an input output sector-country pair and are proxied by patent citations. Our idea is related to trade literature where firms engaging in trade relationships are found to enhance knowledge diffusion (Keller 2004; MacGarvie 2006), but there have been few attempts to study an equivalent effect of knowledge exchange through spillovers. Given our discussion and the ideas we depicted above, we can formulate the following hypotheses:

Hypothesis 1a: Knowledge accumulated in both the input sector-country and in the output sector-country exerts a positive impact on further knowledge flows.

Hypothesis 1b: Knowledge spillovers from external sector-countries (i.e. sector-countries that are not part of the input-output relationship) exert a positive impact on further knowledge flows from the input to the output sector-country.

4.2.2 Technological Distance

Literature on trade flows traditionally focuses on geographical proximity, but in the context of knowledge flows it is important to account for cognitive proximities that are represented by institutional, technological, social and organizational links between economic entities (Paci et al. 2014). In this paper, - beneath geographical distance - we focus on technological distance between industries and countries, the latter being standard in country-level studies on knowledge flows (e.g., Cappelli and Montobbio 2014; Peri 2005). In the context of technological activities, technological distance has been shown to matter most among different other cognitive distances (Paci et al. 2014). Technological distance between countries has been often applied empirically

[4] In the empirical implementation, the external spillovers external to either the input or output sector-country differ only with respect to their weighting scheme (see Section 4.3) as both include the same set of external sector-countries, namely sector-countries that are not part of the input-output relationship.

by using differences in total factor productivity, e.g. between leading countries and laggards (e.g., Aghion et al. 2005). At micro level, it is usually measured with the uncentered correlation between firms' patent portfolios in different technological fields (Jaffe 1986, 1989).

According to Malerba et al. (2013), technological proximity is associated with lower communication and learning costs as firms are better able to recognize and absorb knowledge that is similar to their knowledge base. Hence, the basic expectation is that a larger technological distance between economic entities decreases the probability and the extent of further knowledge flows that might occur between them and can be appropriated by the receiver. More concretely, the larger technological distance is, the less likely further knowledge flows will occur. Consequently, we expect that a larger technological distance between countries, sectors or sector-countries leads to less knowledge flows between them.

Hypothesis 2: Technological distance between countries, sectors and sector-countries has a negative impact on further knowledge flows between an input sector-country and an output sector-country.

4.2.3 Low- vs. High-tech Sectors

Accumulated spillovers from fully external sector-countries might stem from sector-countries that are quite heterogeneous technologically. In addition, input (sector-countries that provide knowledge) and output sector-countries (sector-countries that draw on this knowledge) can vary with respect to their technological orientation and advancement. In innovation and growth literature, high-tech countries or 'technological leaders' are considered as main growth and technological drivers. Mancusi (2008) found that only spillovers from technologically leading countries are effective in increasing innovative output. According to Peri (2005), technologically leading regions may act as learning sources for other regions. Tsai and Wang (2004) found evidence of an R&D spillover effect from the high-tech sector into traditional manufacturing industries in Taiwan. In the same vein, Hu et al. (2005) found that R&D complements technology transfer to developing countries.

In the framework of Cohen and Levinthal (1989, 1990), basic research is thought of broadening a firm's knowledge base and providing it with deeper understanding that is useful for exploiting new technical developments. Success of firms in high-tech sectors is associated with basic research to a larger degree than in low-tech sectors (Czarnitzki and Thorwarth 2012).

Consequently, we expect high-tech sectors to dispose of larger knowledge bases and absorptive capacities. This accumulated knowledge can be a useful source for others, particularly for sector-countries lagging behind the technological frontier. Griffith et al. (2004) found that such laggards catch up particularly fast if they invest heavily in R&D as building an own knowledge stock is necessary to understand knowledge from more advanced countries. The further a country lies behind the frontier, the greater the potential for R&D to increase growth of total factor productivity through technology transfer from more advanced countries. Hence, laggards can benefit disproportionately from high-tech knowledge, but at the same time they need to increase their own absorptive capacity.

Similar to countries or regions, leading high-tech sectors or sector-countries might also act as learning source for others. We argue that knowledge spillovers should only have a positive effect on knowledge flows when the source for spillovers is a high-tech sector-country. We expect that low-tech sector-countries can learn from R&D generated in high-tech sector-countries, but - given the advanced knowledge already accumulated in a high-tech sector-country - the scope for learning for high-tech sector-countries from low-tech sector-countries should be limited. Again, we look at knowledge accumulated within the focal input-output pair and knowledge spilling from other sector-countries, being either high-tech or low-tech. In case of a focal input sector-country being low-tech, we suppose that knowledge from a high-tech output or spillovers from fully external high-tech sector-countries increase further knowledge flows. In case of a low-tech input high-tech output pair, high-tech knowledge that is first learnt by the input sector-country can flow back to the high-tech output sector-country and is utilized there. One can think about the low-tech sector-country being 'fed' with high-tech knowledge, making the low-tech sector-country a more valuable knowledge source for the output sector-country in the aftermath. This idea is summarized in the following hypotheses.

Hypothesis 3a: Knowledge accumulated in the output sector-country only exerts a positive impact on further knowledge flows from the input to the output sector-country if the output sector-country is high-tech.

Hypothesis 3b: Knowledge spillovers from external sector-countries only exert a positive impact on further input-output knowledge flows if the external sectors are high-tech.

Both low-tech and high-tech sector-countries might possess the same need to benefit from high-tech spillovers from the outset. However, we need to take into account that - on the one hand - sector-countries that are farther away from the technological frontier should benefit most since scope for learning is highest there. On the other hand, sector-countries closer to the frontier already have a higher absorptive capacity so that they may be more able to benefit from spillovers (see Aghion et al. (2009) and Miguélez and Moreno (2015), for a similar reasoning). The ease of learning of external knowledge depends on the tradeoff between the relevance of external knowledge and the complexity of this knowledge. Both relevance and complexity are expected to be higher in high-tech sector-countries. Therefore, high-tech spillovers are more relevant, but at the same time more difficult to absorb, especially for low-tech sector-countries. As a consequence, with respect to hypotheses 3a and 3b, we have to leave open the question whether high-tech knowledge spillovers have the potential to generate further knowledge flows only if the input sector-country is high-tech or also if it is low-tech as a low-tech sector-country might lack absorptive capacity to draw on this knowledge.

4.3 The Empirical Model and Estimation

We extend the models proposed by Mancusi (2008) and Peri (2005).[5] Knowledge exchange between input and output sector-countries is assumed to benefit from their knowledge stocks (absorptive capacities) that are given by $I_{ci,si,t}$ and $I_{co,so,t}$ and knowledge spillovers that come from external sector-countries, i.e. sector-countries that are not part of the input-output relationship. ci and co denote the input and output country respectively, si and so the input and output sector and t the year. For the input sector-country, the external spillovers are given by $E_{ci,si,t}$, for the output sector-country by $E_{co,so,t}$. They are proxied with the weighted sum of R&D stocks of sector-countries other than the input and output sector-country. The external stocks need to be weighted appropriately as we cannot assume that external knowledge can be absorbed perfectly. We follow Mancusi (2008) and apply the share of backward citations of sector-country ci, si or co, so in year t as weighting variable denoted as $\phi_{c,s,cj,sj,t}$, i.e. the ci, si or co, so's number of backward citations of an external sector-country cj, sj in ci, si or co, so's total backward citations. Intuitively, the more citations sector-country cj, sj receives from ci, si, the larger the likelihood that its knowledge diffuses to ci, si (Hall et al. 2010, p. 1068). Thus, external spillovers can be defined as follows:

[5] The whole derivations can be found in the Appendix 4.B.1.

$$E_{c,s,t} = \sum_{cj} \sum_{sj} \phi_{c,s,cj,sj} I_{cj,sj,t}, j \neq i, j \neq o \tag{4.3.1}$$

where $c = \{ci; co\}$ and $s = \{si; so\}$, depending on whether we look at spillovers external to ci, si or co, so, $\phi_{c,s,cj,sj,t}$ is the weight and $I_{cj,sj,t}$ the knowledge stock of sector-countries that are external to the input or output sector-country.[6]

We assume that citations are a noisy indicator of actual outflows and estimate a function that depends on the focal sector-countries' knowledge stocks (absorptive capacity), spillovers as defined above, technological distance and other gravity variables $\boldsymbol{x'_{ci,si,co,so}}$ for each input-output pair, industry- and country-specific variables $\boldsymbol{y'_{ci,si,t}}$ and $\boldsymbol{z'_{co,so,t}}$ for both the input and output dimension, and input and output country, industry and time fixed effects.

$$\begin{aligned} C_{ci,si,co,so,t} = exp(&\boldsymbol{x'_{ci,si,co,so,t}}\boldsymbol{\beta} + \boldsymbol{y'_{ci,si,t}}\gamma + \boldsymbol{z'_{co,so,t}}\boldsymbol{\delta} + \alpha_{1i} lnI_{ci,si,t-1} + \alpha_{1o} lnI_{co,so,t-1} + \\ &\alpha_{2i} lnE_{ci,si,t-1} + \alpha_{2o} lnE_{co,so,t-1} + v_{ci} + v_{co} + v_{si} + v_{so} + v_t + \xi_{ci,si,co,so,t}) \end{aligned} \tag{4.3.2}$$

For estimation, we use count data models for the number of forward citations that occur between input and output sector-countries. We estimate both Poisson models and negative binomial models (NBREG), both conditioning on random sector-country pairs (see Boesenberg and Egger 2016).

4.4 Data Sources, Variable Definitions, and Descriptive Statistics

4.4.1 Data Sources

The idea to use input-output tables in order to analyze knowledge flows goes back to Scherer (1984) and has been elaborated on by Verspagen (1997) and Verspagen and De Loo (1999). A technology flow matrix measures how technological knowledge from one sector in a certain country spills over to other sectors in the same or other countries. Patent data comes from the European Patent Office (EPO) PATSTAT database (EPO 2013). We use all published patents between 1995 and 2005 that can be attributed to a technological field according to the Interna-

[6]Please note that input and output sector-countries can be similar. However, to qualify as sector-country that is external to the input-output pair either the country or the sector has to be different.

tional Patent Classification (IPC, WIPO 2014). The number of forward citations is calculated for each year and for each input and output sector-country-pair. In order to avoid truncation of the forward citation counts, we consider 5-year-windows, i.e. forward citations that occur within 5 years after publication of the cited patent (see Squicciarini et al. 2013).
Transformation of technological fields according to IPC subclasses to industries based on aggregated NACE codes is accomplished with concordance data from Lybbert and Zolas (2014). We select the applicant countries and industries according to the World Input Output Tables created by Dietzenbacher et al. (2013).
We match the patent data with country-level data from World Development Indicators (The World Bank 2014) and industry-level data from the OECD STAN and ANBERD databases (OECD 2005, 2012). We end up with 22 countries and 11 industries for 1995 to 2005 that can be used in our estimations. Finally, data on geographical distances and dummies for former colonial ties, common languages and contiguities are assigned (Mayer and Zignago 2011).

4.4.2 Variable Definition

Absorptive Capacity and Spillovers

We construct variables for absorptive capacity for each sector-country and spillovers from external sector-countries. R&D stocks are calculated based on the inventory perpetual method as described in Hall et al. (2010). We use 15% as depreciation rate. As proxy for absorptive capacity, we simply use the one-year lags of the R&D stocks that can be accrued to the focal input and output sector-country pair. For each sector-country pair, we insert the respective R&D stock for both the input and output sector-country as the knowledge flows based on the input sector-country might benefit from both knowledge stocks. We use the one-year lags of the external R&D stocks as proxies for external spillovers that are weighted as already described. The exact definition of $E_{ci,si,t}$ and $E_{co,so,t}$ can be found in table 4.A.1.

Technological Distance

Following the extant literature, we capture technological distance between sector-countries based on the correlation between their share of patents in industries and technologies. The first measure captures differences in technological specialization between two countries and is defined as follows:

$$TECHDIST_{ci,co,t} = 1 - SPECCORR_{ci,co,t} \quad (4.4.1)$$

where $SPECCORR_{ci,co,t}$ is the uncentered correlation coefficient between the share of patents of ci and co in the 17 industries considered here. A value close to 1 indicates a large degree of sectoral specialization.

We essentially expand the existing measure by also including a sectoral dimension, i.e. by measuring whether two sectors or sector-countries are technologically close in specialization. The calculation of distances at this level is involved as both the sectoral assignment of patents and the distance measures are based on IPC (sub)classes. Therefore, we look at each industry separately (the patents were assigned to industries based on an IPC-industry concordance table beforehand), assign all IPC classes that occur in patent applications assigned to a specific sector-country (not only IPC classes that occur in the respective industry definition). Based on this assignment, we calculate distance measures for sector-countries and sectors separately for each sector-country and sector pair. The specialization index becomes $TECHDIST_{ci,si,co,so,t} = 1 - SPECCORR_{ci,si,co,so,t}$ (resp. $TECHDIST_{si,so,t} = 1 - SPECCORR_{si,so,t}$) based on the uncentered correlation between the share of patents of ci, si and co, so (resp. of si and so) in the IPC classes occurring in the underlying patent applications in each year.

Further Variables

The basic specification of a gravity model in the trade literature includes supply factors of the export country, demand factors of the import country, and trade supporting and impeding determinants (geographical and cultural proximity) (Egger and Pfaffermayr 2003). We use the natural logarithm of distance in kilometers between the most populated cities of two countries, denoted by *lndist*, a binary variable measuring if two countries share a land common border, *contig*, and a binary variable, *language*, whether two countries share a common official language as cultural and geographic variables. Finally, we include a binary variable measuring whether former colonial relationships between two countries existed, *colony* (for detailed description of these variables, see table 4.A.1). Following the trade literature, we include both the natural logarithm of GDP and the natural logarithm of GDP per capita as measures of market size and the quality of the economic and institutional environment of a country. Furthermore, we include the percentage of researchers in R&D in a country's population as proxy for a country's human capital. At industry level, we include variables measuring R&D intensity and investment intensity (R&D expenditures and investments as a share of value added) and the natural logarithm of the number of employees as a measure of the size of the industry. Finally, we insert dummy variables that take on value one if the input and output sector, input output country or input output sector-country pairs include the same sectors, countries or sector-countries, respectively, in order to control for the possibility that citations refer heavily to the same country, sector or sector-country.

4.4.3 Descriptive Statistics

Table 4.A.2 shows the summary statistics for the main variables and the sample that is used in the estimations. As expected, the distributions of patent applications and especially of forward citations are very skew. Table 4.A.3 shows summary statistics for some key variables divided into industries and countries. The industries are summarized into a high-tech and low-tech sector based on a OECD definition that relies on R&D intensities (Hatzichronoglou 1997).[7] The high-tech industries are indeed the industries with the highest values for the R&D stock. They also account for an above-average number of patents and forward citations.

4.5 Estimation Results

4.5.1 Basic Results

Input and Output Knowledge Stocks

The basic results for sector-country pairs where the number of forward citations is positive can be found in table 4.A.4.[8] Columns (1), (3) and (5) display coefficients and standard errors from Poisson models, columns (2), (4) and (6) from NBREG. The coefficients for the (weighted)

[7]The same classification will be used later on in the estimations for high-tech and low-tech sectors. To avoid confusion, the unit of observation is the sector-country pair where each sector represents a certain industry. High-tech and low-tech sectors are aggregated sectors and the definition of each comprises several industries as can be seen from Table 4.A.3.

[8]The results from logit models where the incidence of forward citations is the binary dependent variable are shown in 4.B.1.

R&D stocks from both the output and external sector-countries are highly significant and positive in (1), thus suggesting significant spillover effects from external sector-countries on further knowledge flows from the input to the output sector-country. The internal R&D stock's coefficient, however, is not significant in any specification. However, the input sector-country can use knowledge from the output sector-country to increase absorptive capacity on which the output sector-country can build later on. Thus, hypothesis 1a receives support only with respect to knowledge from the output sector-country, hypothesis 1b on the effect of external spillovers receives strong support.

Our measures of technological distance between countries and sector-countries display significantly negative coefficients in all specifications as expected in Hypothesis 2.[9]

With respect to other variables, the number of researchers as a proxy for human capital on both the input and output side strongly increases the number of citations that input patents receive. The same is true for GDP per capita. Most of the time-invariant gravity variables show the expected signs. Notably, the coefficients of technological distance are much larger than those of geographical distance so that geographical distance is found to be of relatively low relevance for international knowledge flows.

External High-tech Spillovers

If we look at external spillovers from high-tech sector-countries, the coefficients get larger, thus suggesting that high-tech sector-countries are valuable spillover sources for input-output knowledge flows (columns (3) and (4)).[10] However, the coefficient of the remainder component (that is the ratio between total external spillovers and high-tech spillovers) is even larger and highly significant so that also non high-tech spillovers can be expected to add to the generation of further knowledge flows relative to high-tech spillovers. In columns (5) and (6) we look at what we call 'top' high-tech spillovers henceforth. These are spillovers from external sector-countries where the sectors belong to the top 10% with respect to R&D intensity.[11] In this case,

[9] The sectoral technological distance shows a positive coefficient in the NBREG, maybe due to correlation with the other distance measures.

[10] The decomposition of the external R&D stock in logarithm in a high-tech and a remainder component is accomplished as follows: We use formula 4.3.2 and look at the following part of the function that is split in a high-tech and a low-tech external stock of knowledge: $ln(E) = ln(E^H + E^L)$. After re-arranging, we get $ln(E) = ln(E^H) + ln(\frac{E^H+E^L}{E^H})$, i.e., the external knowledge stock in logarithm now consists of the high-tech part in logarithm and the ratio between the total external stock and the high-tech stock in logarithm. Both parts are inserted separately in the regression models in order to estimate the extent of the contribution of high-tech external spillovers.

[11] Basically, these are 'Chemicals and Chemical Products' and 'Electrical and Optical Equipment'.

only high-tech spillovers affect knowledge flows significantly. Thus, spillovers from external 'top' high-tech sector-countries are more relevant for the generation of further knowledge flows than high-tech spillovers based on the broader definition. However, the coefficient is smaller than in (3) and (4). In sum, the results yield mixed evidence with respect to hypothesis 3b.

4.5.2 Results for High-tech and Low-tech Sectors

Input and Output Knowledge Stocks

In table 4.A.5, we proceed by looking at knowledge flows occurring between input and output sector-countries being either both low-tech, both high-tech or one being low-tech and the other high-tech.[12] We find that the high-tech input sector-country's knowledge flows are affected by the own knowledge stock (in contrast to the results including the whole sample in Section 4.5.1) if the output sector-country is low-tech (column (3)) and by output knowledge if the output sector-country is also high-tech (1). Spillovers from sector-countries external to the input-sector-country play a role for both input output high-tech and input output low-tech pairs ((1) and (2)), but in (2) spillovers from sector-countries external to the output sector-country do also have an effect. Knowledge flows based on all kind of pairs benefit from the output sector-countries' knowledge stock except for input high-tech output low-tech pairs as then the input sector will avoid drawing on the low-tech knowledge stock provided by the output sector-country.

Knowledge flows based on low-tech input sector-countries benefit disproportionately from knowledge from the output sector-country if the output sector-country is high-tech (4). This indicates that the low-tech sector does better in absorbing high-tech knowledge that is familiar from prior knowledge exchange with the respective output sector-country as compared to absorbing fully external knowledge. The result is perfectly in line with hypothesis 3a. Sector-countries lagging behind the technological frontier not only catch up by learning from more advanced sector-countries, thereby enriching their own knowledge base, they also create the potential for further knowledge flows by drawing on advanced high-tech knowledge spillovers. High-tech sector-countries are thus a source of knowledge for less advanced ones. The pair considered in (4) is

[12] We display results from Poisson estimations only. The main reason lies in the fact that the NBREG yields some implausible coefficients as can be seen from table 4.A.4. Nevertheless, the NBREG results for high-tech vs. low-tech sectors are much more robust compared to the NBREG baseline results in table 4.A.4. Poisson estimates are consistent even if the data is not Poisson distributed provided that the conditional mean is correctly specified. However, one has to use panel-robust standard errors (see Cameron and Trivedi 2005).

also the largest profiteer of spillovers external to the high-tech output sector-country. In sum, low-tech knowledge seems to benefit if combined with a broad array of other sources.

In columns (5) to (8) we use the alternative definition of 'top' high-tech sectors as described in 4.5.1. Some of the associations get stronger indicating on the one hand that 'top' high-tech knowledge is more valuable as knowledge source ((5), (7), (8)), but that this knowledge is also more difficult to draw on (see the smaller coefficients of external spillovers in (6)) for low-tech sector-countries on the other hand.

External High-tech and Low-tech Knowledge Spillovers

In table 4.A.6, we again look at external high-tech spillovers as already done in 4.5.1, but now we also account for the high-tech resp. low-tech sector affiliation of the input and output sector-country as in 4.5.2 above. Knowledge flows benefit from external high-tech knowledge irrespective of which input and output sector-countries are involved, but usually the effect of total spillovers relative to high-tech spillovers is also significant (columns (1)-(4)).

Looking at 'top' high-tech sector-countries and spillovers originating from there, the associations again get stronger for the input and output knowledge stocks. However, external 'top' high-tech spillovers seem only to matter when the input sector-country is low-tech ((6) and (8)). In contrast, the input 'top' sector-countries solely draw on input (7) or output (5) high-tech knowledge.

From the outset, it was not clear whether low-tech sector-countries have enough absorptive capacity in order to absorb more advanced knowledge. The results indicate that knowledge flows can be generated based on low-tech sector-countries if combined with output high-tech knowledge or by combining a wide array of external sources (that are associated with the output sector-country though). Although it seems to be more convenient to draw on output high-tech sector-countries' knowledge stocks directly, we found some evidence that knowledge generation based on low-tech sector-countries also benefit from external 'top' knowledge sources.

In sum, external high-tech spillovers only have the potential to generate further knowledge flows in some cases, namely in combination with other external sources or if constrained to 'top' sector-countries as originators. For low-tech input sector-countries, drawing on knowledge from either an output or external high-tech sectors can generate a broader knowledge base and stimulate own knowledge activities, but the association with the output knowledge stock is generally much stronger.

4.5.3 Results for Different Industries

Input and Output Knowledge Stocks and External Knowledge Spillovers

In this section, we gather further insights into industry specificities with respect to knowledge generation and absorption. To this end, we run the estimations for each input industry separately, not summarizing single industries into high-tech or low-tech sectors. Tables 4.A.7a and 7b show the results for selected low-tech industries and high-tech input industries. Knowledge flows originating in most of the low-tech industries seem to hinge on output knowledge and spillovers external to the respective output sector. In two of the low-tech industries considered here, the internal knowledge stock exerts a negative effect on knowledge flows indicating that the low-tech stock creates a barrier for further flows and the sector-country's ability to generate further knowledge flows depends on output and external knowledge. Interestingly, except for 'Electrical and Optical Equipment' (column (b)(3)), all high-tech industries are 'introverted' and external knowledge absorption does not play any role for further knowledge flows.[13] In sum, the results support the notion that external knowledge absorption is particularly important for low-tech sectors as knowledge generated there can benefit from external knowledge. For technological distance, sectoral distance dominates in low-tech industries, whereas country-level distance seems to dominate in high-tech industries.[14] High-tech industries are already operating at the frontier, but there might be still differences across countries creating barriers for knowledge absorption.

In Figures 4.B.1 and 4.B.2 we go into even deeper detail by looking at industry-industry pairs separately. We display significant coefficients at the 10% significance level for selected input industries, all possible output industries and the following variables: knowledge stocks of the output industry, technological distance at sector-country level, spillovers from external industries (input side) and spillovers from external industries (output side).[15] The results suggest that

[13] This finding is not in line with findings from Belenzon (2012) who found that innovation in the Electronics industry is more cumulative so that fully external knowledge is less valuable. Malerba et al. (2013) argue that international intrasectoral knowledge spillovers are particularly relevant for the Electronics industry that is globalized to a large degree, but in contrast to us they found that national, inter-sectoral knowledge spillovers are relevant for the Chemical industry.

[14] However, geographic distance is not relevant in high-tech industries.

[15] For example, the results in Figure 4.B.1 for input external spillovers should be read as follows: 'If the input industry is 'Agriculture (...)' and the output industry (12) 'Machinery', we find a significantly positive spillover effect from external high-tech industries (i.e., industries that are not 'Agriculture (...)' or 'Machinery') on further

(12) 'Machinery' is the most important high-tech output knowledge source for low-tech industries. In addition, for low-tech industries' knowledge flows, the coefficients of output high-tech knowledge stocks are usually larger in magnitude than output low-tech sources. For high-tech industries, relationships with the output industry (and also with other high-tech industries) seem to be of lower relevance for the generation of knowledge flows. Knowledge flows based on the 'Electrical and Optical Equipment' industry benefit most from output knowledge when the output industry is the same. Interestingly, most of the low-tech industries' knowledge flows are increased by additional external input spillovers even if the output industry is high-tech, suggesting that recombination of knowledge depends on a large variety of sources. In some cases, high-tech industries also complement their knowledge base with appropriate external high-tech or even low-tech knowledge, but the coefficients are generally much smaller than for low-tech industries.

External High-tech and Low-tech Knowledge Spillovers

We again distinguish between external high-tech and low-tech knowledge spillovers (see Tables 4.A.8a – 9b). For three out of four low-tech industries, the spillovers seem to be attributable to 'top' high-tech knowledge from sector-countries external to the output sector. 'Electrical and Optical Equipment' is the only high-tech industry where knowledge flows generally benefit from spillovers from other high-tech industries rather than from other industries.

4.5.4 Robustness of Results

We provide two additional robustness checks.[16] First, we check whether the results are driven by sector-country pairs that consist of the same input and output sector, country or sector-country although we already control for these pairs with dummies. In sum, the results are not affected if we only include sector-countries, sectors or countries that are different from each other.
Second, we check whether the inclusion of "patent scope" and "number of claims" change our results. The patent scope is the technological breadth of a patent measured by the number of technological fields that a patent comprises. The number of claims refer to the legal claims that a patent makes. Both indicators are associated with value and quality of a patent (see

knowledge flows between those industries. In addition, we find a positive association between the 'Machinery' output knowledge stock and further knowledge flows between this industry pair.'

[16] Results are not shown.

Squicciarini et al. 2013). Hence, they might be drivers of forward citations. We check whether the inclusion of industry averages makes other results obsolete. Although their coefficients are highly significant, the other results are not affected at all.

4.6 Conclusions

In this paper, we look at international and inter-sectoral knowledge flows between sector-countries as measured with patent citations. We use a very detailed dataset that goes down to the sector-country input output level. We add to literature by analyzing the effect of absorptive capacity and knowledge spillovers on further knowledge flows that occur between a sector-country pair, by considering technological distance at the sectoral, sector-country and country level, and by providing a detailed analysis for sectors of different technological advancement. This analysis helps to partly resolve the heterogeneity involved in the process of knowledge absorption and generation in an international and inter-sectoral context and to shed light on informal knowledge relationships between technologically diverse input output sector-countries.
The basic estimations show that technological distance, knowledge from output and spillovers from external sector-countries are important drivers of further knowledge flows between input and output sector-countries, but that the output sector-country is very often the most important source of knowledge. Knowledge exchange between input-output sector-countries seems to be mainly a self-sustaining process at first sight where the output sector-country draws on knowledge recombined by the input sector-country but provided by the output sector beforehand. External knowledge spillovers play an important but more limited role in adding to the own knowledge base and generating the basis for further knowledge flows to the output sector. In general, knowledge accumulated in the output sector and external spillovers from sector-countries external to the input or output sector-countries turn out to be more important than the internal knowledge stock of the input sector-country for follow-up knowledge flows.

Estimating the models for input and output sector-countries with different high-tech and low-tech sector affiliations shows that the absorption and utilization of external knowledge spillovers vary across sectors. First, for knowledge flows originating in a low-tech sector-country, drawing on a variety of external knowledge sources (and especially knowledge from the technological frontier) is relatively important. Second, knowledge flows originating in a high-tech sector-country benefit less from enrichment with external high-tech knowledge and absolutely not from external low-tech spillovers. The process of knowledge absorption and generation in high-tech industries seem to take place more in isolation – although high-tech industries have a larger potential to absorb and utilize knowledge from various sources given their high absorptive capacity –, whereas in low-tech industries, learning from a variety of knowledge sources and especially from 'top' high-tech knowledge is prevalent. Third, knowledge flows based on a low-tech sector-country can benefit most from high-tech knowledge coming directly from the output sector-country. This channel might provide more familiar knowledge than drawing on unspecific spillovers. A high-tech high-tech input-output pair is the only constellation where the impact of the output knowledge stock is even larger in magnitude. The knowledge that is first learnt and then recombined by low-tech sector-countries is only valuable for high-tech sector-countries if it is enriched with more advanced knowledge.

Convergence of knowledge across industries that might lead to new, disruptive technologies and products is a well-known phenomenon but its determinants and impacts are poorly understood. The results for high-tech and low-tech sector-countries show that it is mainly established knowledge relationships that thrive further knowledge generation and absorption. Integrating external knowledge is of relatively low importance when industries with more advanced technologies are the direct or indirect receivers of spillovers.

From a policy point of view, the finding that the knowledge flows from most of the high-tech industries do not depend on external knowledge is striking. Obviously, for these industries, only highly specialized knowledge is relevant that might not be available from the external sources considered here. The question arises whether sequential innovation performance in these sector-countries could benefit from more knowledge exchange with external sector-countries and also how low-tech sector-countries could be supported in adopting and using knowledge from the technological frontier.

The major limitation of this study lies in the fact that we use patent citations as proxy for knowledge flows. Although we apply common measures used in literature, the well-known limitations of patent data apply. Unfortunately, data embodying knowledge exchange through labor turnover or migration of knowledge workers is only available in very limited contexts and difficult to obtain at our level of analysis. A further limitation is that we are only able to trace a very limited part of the sequential innovation process. Future research might try to take into account the sequentiality and complexities of knowledge flows and the underlying processes using appropriate methods. It would also help to further refine the empirical analysis by providing firm-level evidence and examining productivity effects of knowledge flows that are based on previous knowledge accumulation and spillovers. Finally, a well-developed theoretical framework for the impact of absorptive capacity and spillovers on further knowledge flows would help understand the underlying processes and interdependencies.

Appendix

4.A Tables

Table 4.A.1: Description of Variables

Variables	Description
C	Number of forward citations within 5 years after publication of patent applications in input sector-country ci, si by output sector-country so, co in year t
l_lnI_i	R&D stock of ci, si in $t-1$ in natural logarithm
l_lnI_o	R&D stock of co, so in $t-1$ in natural logarithm
l_lnE_i	Weighted sum of R&D stocks of $cj, sj, t-1$, $(cj, sj) \neq (co, so), (cj, sj) \neq (ci, si)$, weighted with relative backward citations from ci, si to cj, sj in natural logarithm
l_lnE_o	Weighted sum of R&D stocks of $cj, sj, t-1$, $(cj, sj) \neq (co, so), (cj, sj) \neq (ci, si)$, weighted with relative backward citations from co, so to cj, sj in natural logarithm (Relative backward citations are the number of backward citations from $c, s, t-1$ to $cj, sj, t-T$ divided by the total number of backward citations of $c, s, t-1$ where $T > 1$ and $s = \{si; so\}$ and $c = \{ci; co\}$).
rdint	R&D intensity in ci, si, t and co, so, t
invint	Investment intensity in ci, si, t and co, so, t
lnempln	Natural logarithm of number of employees in ci, si, t and co, so, t
researcher	Researchers in R&D in % in ci, t and co, t
gdppc	GDP per capita in ci, t and co, t
lngdp	Natural logarithm of GDP in ci, t and co, t
techdist_c	Technological distance between ci, t and co, t (uncentered correlation coefficient between the share of patents of ci, t and co, t in 17 industries)
techdist_s	Technological distance between si, t and so, t (uncentered correlation coefficient between the share of patents of si, t and so, t in the underlying technological fields)
techdist_cs	Technological distance between ci, si, t and co, so, t (uncentered correlation coefficient between the share of patents of ci, si, t and co, so, t in the underlying technological fields)
lndist	Geographic distance between ci and co in natural logarithm
contiguity	Dummy for contiguity of ci and co
comlang_off	Dummy for common language of ci and co
colony	Dummy for former colonial relationship between ci and co
c_pair	Dummy indicating whether $ci = co$
s_pair	Dummy indicating whether $si = so$
cs_pair	Dummy indicating whether $ci = co$ and $si = so$

Table 4.A.2: Summary Statistics for Sector-countries with Positive Number of C

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) p50	(6) max
C_t	399,476	62.33	3,161.37	0.00	0.47	876,566.94
lnI_i_t	398,815	19.96	2.68	0.00	20.00	26.49
lnI_o_t	398,277.00	19.86	2.76	0.00	19.94	26.49
lnE_i_t	398,815	22.22	5.44	0.00	23.52	25.46
lnE_o_t	399,418.00	22.26	5.35	0.00	23.51	25.46
$techdist\_s_t$	399,476	0.64	0.26	0.00	0.70	0.98
$techdist\_c_t$	399,476	0.10	0.09	0.00	0.08	0.47
$techdist\_cs_t$	399,476	0.70	0.24	0.00	0.77	1.00
$rdint\_i_t$	399,476	3.47	5.95	0.00	1.23	48.35
$invint\_i_t$	399,476	19.34	12.85	2.30	16.48	210.08
$lnempn\_i_t$	399,476	11.90	1.48	4.48	11.98	16.06
$lngdp\_i_t$	399,476	27.25	1.28	24.04	27.44	30.20
$lngdppc\_i_t$	399,476	10.11	0.31	8.99	10.16	10.70
$researchers\_i_t$	399,476	0.29	0.14	0.10	0.28	0.80
$lndist$	399,476	7.37	1.32	3.98	7.28	9.81

Table 4.A.3: Summary Statistics per Input Sector and Country: Number of Forward Citations, Number of Patents, R&D Stock in Natural Logarithm

	C		$patent\_count_{si,t}$		$lnI_{ci,si,t}$	
si	Mean	Sd	Mean	Sd	Mean	Sd
Low-tech sectors						
Agriculture, Hunting, Forestry and Fishing	23.1	218.2	5,380.2	7,920.9	17.9	3.2
Mining and Quarrying	21.2	185.8	5,575.2	8,894.2	17.6	3.5
Food, Beverages and Tobacco	36.9	708.5	5,766.7	9,792.8	20.2	1.6
Textiles and Textile Products and Leather, Leather and Footwear	36.2	587.2	8,739.2	12,338.6	19.0	1.3
Wood and Products of Wood and Cork	3.2	30.8	859.0	1,149.2	17.4	2.2
Pulp, Paper, Printing and Publishing	33.4	594.5	4,468.0	6,864.8	19.2	1.8
Coke, Refined Petroleum and Nuclear Fuel	16.6	230.5	1,928.1	2,765.7	20.0	2.0
Rubber and Plastics	13.0	139.0	1,849.3	2,765.4	20.2	1.5
Other Non-Metallic Mineral	17.7	201.1	2,454.5	3,521.5	19.4	1.8
Basic Metals and Fabricated Metal	69.8	1,179.6	9,530.4	15,449.3	20.5	1.6
Manufacturing, nec; Recycling	9.1	115.7	1,623.2	2,466.4	18.8	2.6
Electricity, Gas and Water Supply	14.3	157.6	2,427.5	3,503.8	19.1	3.1
Construction	28.2	316.7	5,249.6	7,111.6	18.8	2.2
High-tech sectors						
Chemicals and Chemical Products	132.9	2,243.1	18,535.0	35,287.1	22.2	1.8
Machinery, nec	83.9	1,444.9	12,152.6	20,252.3	21.2	1.7
Electrical and Optical Equipment	365.8	11,580.8	31,308.1	59,384.2	22.9	1.8
Transport Equipment	46.0	545.3	6,575.0	9,053.2	22.4	2.4
Total	30.3	2,209.5	4,907.8	16,457.9	19.0	3.3
	C		$patent\_count_{ci,t}$		$lnI_{ci,si,t}$	
ci	Mean	Sd	Mean	Sd	Mean	Sd
AT	5.1	26.3	10,682.0	458.5	19.2	1.7
AU	2.5	8.3	8,508.6	2,073.7	19.3	0.5
BE	7.4	52.3	8,179.6	1,013.1	19.6	1.6
CA	27.3	485.5	21,500.2	3,652.9	20.5	1.3
CZ	0.7	2.9	2,448.0	169.8	18.3	1.8
DE	100.7	688.6	188,960.2	15,506.0	21.5	1.8
ES	3.4	20.2	13,742.9	1,376.6	20.0	1.3
FI	5.9	37.1	16,177.0	1,242.0	18.8	1.3
FR	28.8	253.2	70,182.6	5,204.7	21.5	1.6
GB	29.7	326.2	48,142.0	2,936.2	21.5	1.6
GR	0.4	1.9	344.7	37.7	17.4	1.0
HU	0.8	5.1	2,407.1	284.8	16.8	2.6
IE	2.7	26.8	2,701.6	348.4	17.7	2.3
IT	8.8	57.4	25,401.3	1,969.7	19.9	2.9
KR	43.8	731.4	104,958.3	51,878.2	20.6	1.8
NL	22.5	220.1	30,191.2	6,536.6	19.8	1.6
PL	0.8	5.1	5,084.3	498.2	18.8	1.2
PT	0.4	2.0	526.4	111.9	17.4	1.2
SI	0.5	2.0	611.5	140.8	16.4	2.8
SK	0.5	2.2	436.6	52.2	11.9	8.2
US	609.9	12,256.9	410,550.0	57,957.4	23.3	1.7
Total	30.3	2,209.5	45,547.7	88,706.3	19.0	3.3

Table 4.A.4: Count-data Models, Dependent Variable: Number of Forward Citations that Input Sector-country Receives from Output Sector-country

VARIABLES	(1) Poisson fw_cit5	(2) NBREG fw_cit5	(3) Poisson fw_cit5	(4) NBREG fw_cit5	(5) Poisson fw_cit5	(6) NBREG fw_cit5
l_lnI_i	0.040	-0.003	0.018	-0.003	0.046	-0.003
	(0.058)	(0.003)	(0.058)	(0.003)	(0.061)	(0.003)
l_lnI_o	0.151***	0.003	0.141***	0.003	0.158***	0.003
	(0.043)	(0.003)	(0.041)	(0.003)	(0.049)	(0.003)
l_lnE_i	0.013***	0.009***				
	(0.003)	(0.001)				
l_lnE_o	0.013***	0.009***				
	(0.003)	(0.001)				
l_lnE_high_i			0.035***	0.014***		
			(0.008)	(0.001)		
l_lnE_high_ratio_i			0.974***	0.225***		
			(0.340)	(0.029)		
l_lnE_high_o			0.013	0.012***		
			(0.011)	(0.001)		
l_lnE_high_ratio_o			0.015	0.110***		
			(0.455)	(0.019)		
l_lnE_high_top_i					0.014***	0.009***
					(0.003)	(0.001)
l_lnE_high_top_ratio_i					-0.003	0.003
					(0.043)	(0.003)
l_lnE_high_top_o					0.014***	0.009***
					(0.003)	(0.001)
l_lnE_high_top_ratio_o					-0.037	0.008***
					(0.056)	(0.003)
rdint_i	0.008	0.002***	0.006	0.002***	0.008	0.002***
	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)
rdint_o	0.001	0.001**	0.000	0.001**	0.001	0.002**
	(0.008)	(0.001)	(0.008)	(0.001)	(0.008)	(0.001)
invint_i	0.001	-0.001***	0.000	-0.001***	0.001	-0.001***
	(0.003)	(0.000)	(0.003)	(0.000)	(0.003)	(0.000)
invint_o	-0.001	-0.001***	-0.002	-0.001***	-0.001	-0.001***
	(0.003)	(0.000)	(0.003)	(0.000)	(0.003)	(0.000)
lnempn_i	-0.108	0.057***	-0.003	0.058***	-0.127	0.057***
	(0.160)	(0.009)	(0.167)	(0.009)	(0.188)	(0.009)
lnempn_o	-0.158	0.068***	-0.123	0.068***	-0.191	0.068***
	(0.169)	(0.009)	(0.171)	(0.009)	(0.202)	(0.009)
researchers_i	2.517***	0.787***	2.519***	0.786***	2.516***	0.787***
	(0.514)	(0.075)	(0.504)	(0.075)	(0.515)	(0.075)
researchers_o	1.834***	0.596***	1.825***	0.591***	1.839***	0.596***
	(0.436)	(0.075)	(0.429)	(0.075)	(0.437)	(0.075)
lngdppc_i	7.117***	1.950***	7.668***	2.002***	7.016***	1.942***
	(0.929)	(0.165)	(0.936)	(0.165)	(0.959)	(0.165))
lngdppc_o	6.338***	1.342***	6.427***	1.372***	6.053***	1.339***
	(0.837)	(0.161)	(0.837)	(0.161)	(0.870)	(0.161)
lngdp_i	-6.091***	-1.641***	-6.624***	-1.691***	-5.990***	-1.634***
	(0.823)	(0.145)	(0.825)	(0.146)	(0.810)	(0.146)
lngdp_o	-4.882***	-0.741***	-4.998***	-0.768***	-4.616***	-0.737***
	(0.797)	(0.142)	(0.813)	(0.142)	(0.752)	(0.142)
techdist_s	-0.772	0.122***	-0.699	0.117***	-0.693	0.134***
	(0.536)	(0.031)	(0.458)	(0.031)	(0.616)	(0.031)
techdist_cs	-1.394***	-0.490***	-1.392***	-0.491***	-1.390***	-0.490***
	(0.292)	(0.028)	(0.282)	(0.028)	(0.292)	(0.028)
techdist_c	-2.232***	-1.050***	-2.116***	-1.049***	-2.238***	-1.051***
	(0.678)	(0.052)	(0.672)	(0.052)	(0.680)	(0.052)
lndist	-0.170***	0.032***	-0.174***	0.032***	-0.169***	0.032***
	(0.018)	(0.006)	(0.018)	(0.006)	(0.018)	(0.006)
contig	0.166***	0.206***	0.167***	0.206***	0.167***	0.206***
	(0.023)	(0.016)	(0.023)	(0.016)	(0.024)	(0.016)
comlang_off	0.171***	0.153***	0.169***	0.153***	0.174***	0.152***
	(0.021)	(0.016)	(0.021)	(0.016)	(0.022)	(0.016)
colony	0.059***	-0.005	0.061***	-0.004	0.057**	-0.004
	(0.022)	(0.017)	(0.022)	(0.017)	(0.024)	(0.017)
c_pair	2.338***	0.378***	2.341***	0.378***	2.349***	0.378***
	(0.054)	(0.019)	(0.053)	(0.019)	(0.058)	(0.019)
s_pair	0.680**	0.106***	0.738***	0.107***	0.715**	0.111***
	(0.315)	(0.018)	(0.272)	(0.018)	(0.357)	(0.018)
cs_pair	0.112	-0.072**	0.088	-0.074**	0.125	-0.071** *
	(0.102)	(0.033)	(0.086)	(0.033)	(0.115)	(0.033)
Observations	399,476	399,476	399,476	399,476	399,476	399,476
Log-likelihood	-2.390e+06	-751758	-2.384e+06	-751716	-2.388e+06	-751750
Wald chi2	270148	59715	272871	59849	260956	59742

(Cluster robust) standard errors (for the Poisson) in parentheses, constants are suppressed.
Time, input sector, output sector, input country and output country dummies are always included.
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.5: Poisson Models for High-tech and Low-tech Sectors, Dependent Variable: Number of Forward Citations that Input Sector-country Receives from Output Sector-country

VARIABLES	(1) Input high-tech & output high-tech	(2) Input low-tech & output low-tech	(3) Input high-tech & output low-tech	(4) Input low-tech & output high-tech	(5) Input 'top' high-tech & output 'top' high-tech	(6) Input low-tech & output low-tech	(7) Input 'top' high-tech & output low-tech	(8) Input low-tech & output 'top' high-tech
l_lnI_i	-0.026	-0.005	0.494***	-0.079	0.221	-0.028	1.071***	-0.083
	(0.316)	(0.022)	(0.161)	(0.061)	(0.499)	(0.028)	(0.301)	(0.086)
l_lnI_o	0.873***	0.032*	-0.051	0.522***	1.552***	0.063**	-0.041	0.640***
	(0.243)	(0.018)	(0.063)	(0.142)	(0.418)	(0.026)	(0.074)	(0.234)
l_lnE_i	0.013***	0.019**	0.010	0.002	0.005	0.009**	0.016	0.002
	(0.005)	(0.007)	(0.009)	(0.008)	(0.008)	(0.004)	(0.010)	(0.008)
l_lnE_o	0.003	0.017***	0.007	0.022***	0.002	0.016***	0.004	0.025***
	(0.004)	(0.005)	(0.006)	(0.008)	(0.006)	(0.004)	(0.005)	(0.009)
rdint_i	0.007	0.071**	-0.012	0.088***	-0.006	-0.002	-0.015	-0.009
	(0.008)	(0.030)	(0.014)	(0.032)	(0.005)	(0.013)	(0.014)	(0.022)
rdint_o	-0.010	0.055*	0.095*	0.002	-0.006	-0.007	0.026	-0.004
	(0.007)	(0.029)	(0.057)	(0.014)	(0.005)	(0.010)	(0.032)	(0.015)
invint_i	0.002	0.001	0.017**	0.002	0.013**	0.002	0.013*	0.002
	(0.006)	(0.002)	(0.008)	(0.002)	(0.006)	(0.002)	(0.008)	(0.003)
invint_o	0.005	-0.002	0.002	0.001	-0.001	0.001	0.003	-0.003
	(0.006)	(0.002)	(0.003)	(0.007)	(0.007)	(0.003)	(0.003)	(0.006)
lnempn_i	-0.346	0.104	0.054	0.120	-0.365	0.095	-0.427	0.408**
	(0.390)	(0.094)	(0.303)	(0.146)	(0.485)	(0.104)	(0.456)	(0.184)
lnempn_o	-0.494	-0.015	-0.058	0.002	-0.875*	0.003	0.126	0.106
	(0.414)	(0.090)	(0.182)	(0.201)	(0.531)	(0.103)	(0.279)	(0.243)
researchers_i	2.408***	2.881***	2.070***	1.464*	0.803	2.625***	0.893	1.124
	(0.819)	(0.494)	(0.787)	(0.799)	(1.228)	(0.366)	(1.025)	(0.765)
researchers_o	0.153	2.247***	1.132*	1.363**	-1.426	1.930***	0.735	0.990
	(0.837)	(0.475)	(0.583)	(0.620)	(1.377)	(0.358)	(0.685)	(0.687)
lngdppc_i	8.655***	6.456***	4.535**	8.562***	10.532***	6.403***	8.224***	9.671***
	(2.410)	(1.068)	(2.245)	(1.264)	(3.701)	(0.833)	(2.681)	(1.495)
lngdppc_o	7.491***	8.071***	5.249***	5.949***	10.733***	7.834***	4.872***	7.337***
	(2.307)	(1.087)	(1.580)	(1.372)	(3.683)	(0.804)	(1.680)	(1.630)
lngdp_i	-7.593***	-5.363***	-4.000**	-6.600***	-10.187***	-4.974***	-7.889***	-7.667***
	(2.241)	(0.971)	(1.853)	(1.295)	(3.693)	(0.762)	(2.220)	(1.393)
lngdp_o	-6.237***	-6.486***	-4.200***	-4.841***	-9.563***	-6.197***	-3.933***	-6.232***
	(2.273)	(1.000)	(1.358)	(1.160)	(3.697)	(0.728)	(1.401)	(1.339)
techdist_s	0.334	-1.640***	-0.300	-1.887***	-8.623	-1.267***	-0.183	-1.191*
	(2.387)	(0.280)	(0.774)	(0.553)	(6.395)	(0.359)	(0.946)	(0.723)
techdist_cs	-1.817**	-1.349***	-1.364***	-0.672*	0.536	-1.463***	-1.355**	-1.480***
	(0.872)	(0.169)	(0.455)	(0.399)	(1.346)	(0.203)	(0.571)	(0.526)
techdist_c	-3.468***	-1.428***	-1.198*	-0.236	-3.793**	-0.790***	-1.404*	-0.250
	(1.253)	(0.301)	(0.633)	(0.517)	(1.502)	(0.278)	(0.792)	(0.658)
lndist	-0.155***	-0.200***	-0.203***	-0.211***	-0.143**	-0.205***	-0.202***	-0.211***
	(0.045)	(0.012)	(0.025)	(0.023)	(0.072)	(0.011)	(0.031)	(0.028)
contig	0.146	0.123***	0.131**	0.203***	0.009	0.159***	0.127*	0.206***
	(0.091)	(0.022)	(0.052)	(0.050)	(0.119)	(0.021)	(0.067)	(0.068)
comlang_off	0.173**	0.168***	0.234***	0.193***	0.262***	0.168***	0.209***	0.164***
	(0.072)	(0.023)	(0.042)	(0.041)	(0.101)	(0.020)	(0.052)	(0.053)
colony	0.028	0.077***	0.076*	0.086**	0.056	0.083***	0.132**	0.148***
	(0.072)	(0.024)	(0.044)	(0.041)	(0.115)	(0.020)	(0.058)	(0.057)
c_pair	2.315***	2.278***	2.422***	2.582***	2.788***	2.382***	2.459***	2.583***
	(0.137)	(0.040)	(0.086)	(0.081)	(0.273)	(0.037)	(0.105)	(0.099)
s_pair	1.412	0.108			-3.675	0.309		
	(1.712)	(0.180)			(5.141)	(0.212)		
cs_pair	0.043	0.153*			-0.272	0.129		
	(0.183)	(0.087)			(0.290)	(0.083)		
Observations	28,697	218,795	75,928	76,056	7,804	345,701	45,971	45,989
Log-likelihood	-598804	-671057	-535082	-503795	-313291	-1.500e+06	-459333	-428626
Wald chi2	57155	143799	71093	75724	47833	243671	56387	56479

Cluster robust standard errors in parentheses, constants are suppressed.
Time, input sector, output sector, input country and output country dummies are always included.
*** p<0.01, ** p<0.05, * p<0.1

Table 4.A.6: Poisson Models for High-tech and Low-tech Sectors with External High-tech and Low-tech R&D Stocks, Dependent Variable: Number of Forward Citations that Input Sector-country Receives from Output Sector-country

VARIABLES	(1) Input high-tech & output high-tech	(2) Input low-tech & output low-tech	(3) Input high-tech & output low-tech	(4) Input low-tech & output high-tech	(5) Input 'top' high-tech & output 'top' high-tech	(6) Input low-tech & output low-tech	(7) Input 'top' high-tech & output low-tech	(8) Input low-tech & output 'top' high-tech
l_lnI_i	-0.083	-0.003	0.362***	-0.080	0.224	-0.028	0.971***	-0.079
	(0.311)	(0.021)	(0.135)	(0.059)	(0.507)	(0.028)	(0.303)	(0.084)
l_lnI_o	0.846***	0.029	-0.053	0.435***	1.565***	0.059**	-0.043	0.705***
	(0.243)	(0.018)	(0.062)	(0.123)	(0.438)	(0.024)	(0.074)	(0.154)
l_lnE_high_i	0.033***	-0.039	0.041**	0.027				
	(0.011)	(0.044)	(0.017)	(0.022)				
l_lnE_high_ratio_i	0.986**	-2.119	1.342**	0.978				
	(0.392)	(1.639)	(0.657)	(0.765)				
l_lnE_high_o	0.002	0.077*	0.057***	0.047***				
	(0.011)	(0.042)	(0.018)	(0.012)				
l_lnE_high_ratio_o	-0.091	2.197	1.953***	1.107**				
	(0.421)	(1.543)	(0.688)	(0.559)				
l_lnE_high_top_i					0.008	0.010**	0.016	0.003
					(0.007)	(0.004)	(0.011)	(0.008)
l_lnE_high_top_ratio_i					-0.033	0.070	0.160	0.251**
					(0.069)	(0.090)	(0.227)	(0.125)
l_lnE_high_top_o					0.007	0.017***	0.005	0.027***
					(0.005)	(0.004)	(0.005)	(0.010)
l_lnE_high_top_ratio_o					-0.081	0.090	0.008	-0.116
					(0.112)	(0.096)	(0.104)	(0.206)
rdint_i	0.006	0.073**	-0.013	0.084**	-0.006	-0.001	-0.015	-0.006
	(0.007)	(0.030)	(0.014)	(0.033)	(0.005)	(0.013)	(0.013)	(0.022)
rdint_o	-0.009	0.052*	0.090	0.001	-0.006	-0.007	0.028	-0.005
	(0.007)	(0.029)	(0.055)	(0.015)	(0.005)	(0.010)	(0.031)	(0.014)
invint_i	0.000	0.001	0.014**	0.002	0.014**	0.002	0.011**	0.002
	(0.006)	(0.002)	(0.007)	(0.002)	(0.006)	(0.002)	(0.005)	(0.003)
invint_o	0.004	-0.002	0.001	-0.002	0.001	0.001	0.003	-0.001
	(0.006)	(0.002)	(0.003)	(0.006)	(0.007)	(0.003)	(0.003)	(0.005)
lnempn_i	-0.132	0.101	0.308	0.147	-0.399	0.098	-0.257	0.439**
	(0.409)	(0.093)	(0.306)	(0.141)	(0.504)	(0.106)	(0.606)	(0.187)
lnempn_o	-0.485	-0.012	-0.023	0.190	-0.934	0.025	0.121	-0.007
	(0.428)	(0.089)	(0.184)	(0.208)	(0.571)	(0.100)	(0.277)	(0.340)
researchers_i	2.496***	2.837***	2.236***	1.432*	0.823	2.615***	1.058	1.010
	(0.830)	(0.493)	(0.767)	(0.800)	(1.198)	(0.367)	(0.912)	(0.789)
researchers_o	0.217	2.287***	1.144**	1.451**	-1.359	1.933***	0.751	0.910
	(0.849)	(0.471)	(0.582)	(0.624)	(1.322)	(0.357)	(0.690)	(0.678)
lngdppc_i	9.004***	6.490***	4.618**	8.876***	10.012***	6.423***	8.826***	9.880***
	(2.370)	(1.064)	(2.305)	(1.292)	(3.570)	(0.829)	(2.500)	(1.466)
lngdppc_o	7.509***	8.086***	5.621***	5.946***	9.561***	7.972***	4.884***	6.619***
	(2.254)	(1.077)	(1.498)	(1.344)	(3.335)	(0.781)	(1.710)	(1.500)
lngdp_i	-7.948***	-5.397***	-4.113**	-6.930***	-9.740***	-4.993***	-8.381***	-7.790***
	(2.181)	(0.967)	(1.914)	(1.346)	(3.562)	(0.761)	(2.105)	(1.388)
lngdp_o	-6.273***	-6.495***	-4.596***	-4.867***	-8.491***	-6.327***	-3.950***	-5.596***
	(2.219)	(0.989)	(1.268)	(1.131)	(3.285)	(0.726)	(1.407)	(1.378)
techdist_s	0.580	-1.633***	0.003	-1.693***	-7.215	-1.225***	-0.197	-1.208*
	(1.929)	(0.279)	(0.808)	(0.578)	(6.519)	(0.349)	(0.909)	(0.716)
techdist_cs	-1.760**	-1.354***	-1.433***	-0.745*	0.532	-1.490***	-1.386**	-1.397***
	(0.838)	(0.169)	(0.431)	(0.401)	(1.361)	(0.201)	(0.553)	(0.520)
techdist_c	-3.378***	-1.431***	-1.001	-0.104	-3.705**	-0.785***	-1.410*	-0.344
	(1.252)	(0.297)	(0.639)	(0.531)	(1.469)	(0.282)	(0.795)	(0.660)
lndist	-0.157***	-0.200***	-0.203***	-0.211***	-0.130*	-0.205***	-0.203***	-0.209***
	(0.046)	(0.012)	(0.025)	(0.023)	(0.077)	(0.011)	(0.031)	(0.028)
contig	0.157*	0.123***	0.141***	0.209***	0.028	0.158***	0.128*	0.199***
	(0.093)	(0.022)	(0.054)	(0.051)	(0.122)	(0.021)	(0.068)	(0.068)
comlang_off	0.171**	0.169***	0.226***	0.188***	0.278**	0.168***	0.209***	0.166***
	(0.072)	(0.023)	(0.042)	(0.042)	(0.108)	(0.020)	(0.052)	(0.052)
colony	0.035	0.076***	0.078*	0.087**	0.039	0.083***	0.134**	0.145***
	(0.073)	(0.024)	(0.045)	(0.042)	(0.129)	(0.020)	(0.058)	(0.056)
c_pair	2.311***	2.278***	2.442***	2.594***	2.918***	2.381***	2.455***	2.581***
	(0.132)	(0.040)	(0.087)	(0.082)	(0.338)	(0.037)	(0.105)	(0.098)
s_pair	1.740	0.109			-5.086	0.319		
	(1.427)	(0.179)			(5.580)	(0.214)		
cs_pair	0.060	0.154*			-0.369	0.126		
	(0.181)	(0.087)			(0.304)	(0.083)		
Observations	28,697	218,795	75,928	76,056	7,804	345,701	45,971	45,989
Log-likelihood	-594991	-670863	-532575	-502553	-311206	-1.499e+06	-458889	-427867
Wald chi2	54412	144216	70210	75482	45327	244367	56184	59197

Cluster robust standard errors in parentheses, constants are suppressed.
Time, input sector, output sector, input country and output country dummies are always included.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

VARIABLES	(1)	(2)	(3)	(4)
l_lnI_i	-0.037	-0.091***	0.351***	-0.125***
	(0.054)	(0.032)	(0.120)	(0.046)
l_lnI_o	0.095*	0.128***	0.286***	0.086**
	(0.052)	(0.039)	(0.092)	(0.041)
l_lnE_i	-0.205	-0.021	0.766***	0.005
	(0.251)	(0.190)	(0.213)	(0.003)
l_lnE_o	0.014**	0.024***	0.022***	0.004
	(0.006)	(0.004)	(0.007)	(0.004)
rdint_i	0.315**	0.146	-0.026	-0.171
	(0.135)	(0.091)	(0.092)	(0.114)
rdint_o	-0.016	-0.013	0.006	-0.012
	(0.011)	(0.011)	(0.013)	(0.011)
invint_i	-0.009	0.003	-0.020	0.027
	(0.013)	(0.006)	(0.013)	(0.027)
invint_o	0.007*	0.004	0.006	0.006
	(0.004)	(0.005)	(0.007)	(0.004)
lnempn_i	-0.275	0.059	2.643***	-0.487**
	(0.318)	(0.103)	(0.360)	(0.229)
lnempn_o	0.496***	0.321***	0.736***	0.140
	(0.174)	(0.079)	(0.246)	(0.202)
researchers_i	-2.028	3.843***	3.029***	4.518***
	(1.287)	(0.381)	(1.132)	(0.661)
researchers_o	0.463	2.382***	2.329**	2.315***
	(0.849)	(0.418)	(1.038)	(0.679)
lngdppc_i	0.713	-1.227***	-0.020	-2.271***
	(0.962)	(0.296)	(0.489)	(0.376)
lngdppc_o	0.532	-0.978***	0.778	-1.046***
	(0.699)	(0.246)	(0.517)	(0.405)
lngdp_i	0.739***	0.890***	-2.184***	1.553***
	(0.272)	(0.089)	(0.425)	(0.235)
lngdp_o	0.201	0.310***	-0.204	0.626***
	(0.177)	(0.092)	(0.303)	(0.207)
techdist_s	-2.058**	-4.404***	-5.256***	-3.584***
	(0.941)	(0.534)	(1.344)	(0.995)
techdist_cs	-0.084	-0.192	0.339	-0.398
	(0.981)	(0.424)	(0.721)	(0.435)
techdist_c	-0.811	-5.380***	-2.301**	-2.804***
	(1.113)	(0.568)	(1.069)	(0.493)
lndist	0.051	-0.437***	-0.268***	-0.136***
	(0.150)	(0.041)	(0.098)	(0.051)
contig	0.886***	-0.332***	0.257	0.635***
	(0.256)	(0.109)	(0.201)	(0.145)
comlang_off	0.328	0.764***	0.573**	0.740***
	(0.225)	(0.104)	(0.247)	(0.094)
colony	-0.730*	-0.271**	-0.780***	-0.260**
	(0.401)	(0.106)	(0.202)	(0.129)
c_pair	3.148***	1.126***	2.543***	2.453***
	(0.422)	(0.143)	(0.366)	(0.199)
s_pair	-0.113	-1.802***	-1.693*	-0.958
	(0.830)	(0.470)	(0.881)	(0.622)
cs_pair	0.227	0.344	0.293	0.084
	(0.353)	(0.343)	(0.638)	(0.326)
Observations	16,404	19,284	22,540	26,428
Log-likelihood	-94056	-22139	-83920	-83439
Wald chi2	5117	3753	1712	8119

Cluster robust standard errors in parentheses,
constants are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

(a) Selected Low-tech Input Industries

VARIABLES	(1)	(2)	(3)	(4)
l_lnI_i	0.254	0.365*	0.594	0.418
	(0.694)	(0.220)	(1.172)	(0.344)
l_lnI_o	0.054	0.049	0.543*	0.114
	(0.074)	(0.099)	(0.318)	(0.098)
l_lnE_i	-0.002	-0.008**	0.020***	0.006
	(0.006)	(0.004)	(0.006)	(0.010)
l_lnE_o	0.008	0.001	0.010*	-0.000
	(0.008)	(0.004)	(0.006)	(0.009)
rdint_i	-0.055***	0.009	0.034***	0.002
	(0.015)	(0.029)	(0.010)	(0.013)
rdint_o	-0.006	-0.031***	0.023*	-0.033***
	(0.008)	(0.008)	(0.014)	(0.012)
invint_i	-0.012*	0.008	-0.004	-0.006
	(0.007)	(0.015)	(0.017)	(0.007)
invint_o	0.005**	0.001	-0.017*	0.005
	(0.003)	(0.006)	(0.009)	(0.006)
lnempn_i	0.992	1.400***	-0.783	-1.130
	(0.994)	(0.337)	(1.807)	(0.870)
lnempn_o	0.329	0.220	-0.059	0.241
	(0.236)	(0.276)	(0.391)	(0.320)
researchers_i	-2.073	2.814***	2.768	4.318***
	(1.752)	(0.551)	(2.887)	(1.328)
researchers_o	2.413***	3.358***	1.573	2.062
	(0.798)	(0.579)	(1.224)	(1.270)
lngdppc_i	1.295	0.681	16.518	-3.131***
	(0.866)	(0.611)	(11.684)	(0.928)
lngdppc_o	0.304	-0.660	6.747**	-0.621
	(0.836)	(0.585)	(3.140)	(0.517)
lngdp_i	-0.634	-1.162***	-16.322	1.618***
	(0.460)	(0.364)	(11.516)	(0.596)
lngdp_o	-0.312	0.122	-5.892*	0.537
	(0.530)	(0.416)	(3.250)	(0.331)
techdist_s	0.885	-2.307*	-1.901	-1.819
	(1.338)	(1.205)	(1.851)	(1.574)
techdist_cs	-2.213***	-2.714***	-0.546	-1.064
	(0.683)	(0.611)	(0.833)	(0.957)
techdist_c	-2.808***	-1.976**	-5.374***	-2.196**
	(0.820)	(0.894)	(1.895)	(1.067)
lndist	0.373	0.334	11.171	-0.071
	(0.262)	(0.271)	(7.598)	(0.177)
contig	1.015*	0.928	7.573	0.779**
	(0.561)	(0.627)	(5.516)	(0.367)
comlang_off	0.336	1.001***	9.668	0.537***
	(0.303)	(0.186)	(7.073)	(0.168)
colony	0.841	0.570	-4.648	-0.318
	(0.686)	(0.467)	(5.968)	(0.221)
c_pair	3.463***	3.461***	24.324**	2.517***
	(0.754)	(0.728)	(12.167)	(0.389)
s_pair	2.142***	-0.945*	-0.619	-0.238
	(0.668)	(0.544)	(2.062)	(1.251)
cs_pair	0.052	-0.151	0.414	0.195
	(0.486)	(0.388)	(1.054)	(0.434)
Observations	26,100	29,273	27,675	21,577
Log-likelihood	-350988	-176745	-519514	-144477
Wald chi2	9930	11801	7120	6108

Cluster robust standard errors in parentheses,
constants are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

(b) High-tech Input Industries

Table 4.A.7: Poisson Models for Different Input Industries, Dependent Variable: Number of Forward Citations that Input Sector-country Receives from Output Sector-country

The column numbers refer to the following input industries: (a)(1) Agriculture, Hunting, Forestry and Fishing (2) Wood and Products of Wood and Cork (3) Pulp, Paper, Printing and Publishing (4) Construction; (b)(1) Chemicals and Chemical Products (2) Machinery, nec (3) Electrical and Optical Equipment (4) Transport Equipment.

VARIABLES	(1)	(2)	(3)	(4)
l_lnI_i	-0.032	-0.078***	0.279***	-0.120***
	(0.048)	(0.026)	(0.099)	(0.046)
l_lnI_o	0.090*	0.119***	0.265***	0.069
	(0.054)	(0.031)	(0.088)	(0.043)
l_lnE_high_i	-0.254	0.043	1.070***	0.101***
	(0.275)	(0.228)	(0.201)	(0.039)
l_lnE_high_ratio_i	-1.974	0.364	13.502***	3.657**
	(2.502)	(0.897)	(2.994)	(1.454)
l_lnE_high_o	0.036	0.112***	0.046	0.010
	(0.023)	(0.019)	(0.051)	(0.026)
l_lnE_high_ratio_o	1.029	3.311***	0.930	0.234
	(1.076)	(0.703)	(1.972)	(0.960)
rdint_i	0.339**	0.136	-0.039	-0.173
	(0.148)	(0.091)	(0.089)	(0.108)
rdint_o	-0.017	-0.018*	0.008	-0.015
	(0.011)	(0.011)	(0.013)	(0.010)
invint_i	-0.009	0.001	-0.022*	0.026
	(0.013)	(0.006)	(0.013)	(0.027)
invint_o	0.006	0.003	0.007	0.006
	(0.005)	(0.005)	(0.007)	(0.004)
lnempn_i	-0.283	0.071	2.675***	-0.513**
	(0.336)	(0.104)	(0.369)	(0.229)
lnempn_o	0.510***	0.347***	0.635***	0.188
	(0.179)	(0.078)	(0.178)	(0.194)
researchers_i	-2.040	3.672***	3.653***	4.520***
	(1.317)	(0.377)	(1.102)	(0.652)
researchers_o	0.489	2.435***	2.223**	2.234***
	(0.849)	(0.394)	(0.964)	(0.697)
lngdppc_i	0.753	-1.274***	-0.048	-2.297***
	(0.964)	(0.322)	(0.490)	(0.370)
lngdppc_o	0.611	-0.972***	0.710	-0.973**
	(0.716)	(0.245)	(0.494)	(0.393)
lngdp_i	0.737***	0.864***	-2.152***	1.580***
	(0.277)	(0.089)	(0.420)	(0.234)
lngdp_o	0.182	0.306***	-0.075	0.592***
	(0.186)	(0.089)	(0.237)	(0.198)
techdist_s	-2.129**	-4.375***	-5.173***	-3.533***
	(0.918)	(0.531)	(1.379)	(0.973)
techdist_cs	-0.057	-0.204	0.356	-0.366
	(1.027)	(0.416)	(0.758)	(0.421)
techdist_c	-0.853	-5.342***	-2.181**	-2.658***
	(1.170)	(0.535)	(1.086)	(0.503)
lndist	0.069	-0.433***	-0.264***	-0.136***
	(0.159)	(0.041)	(0.090)	(0.050)
contig	0.915***	-0.334***	0.255	0.649***
	(0.259)	(0.106)	(0.184)	(0.142)
comlang_off	0.311	0.776***	0.565**	0.742***
	(0.229)	(0.105)	(0.232)	(0.092)
colony	-0.731*	-0.282***	-0.809***	-0.256**
	(0.400)	(0.105)	(0.200)	(0.125)
c_pair	3.193***	1.141***	2.525***	2.475***
	(0.439)	(0.141)	(0.316)	(0.196)
s_pair	-0.129	-2.014***	-1.475	-0.998*
	(0.796)	(0.471)	(0.904)	(0.605)
cs_pair	0.234	0.261	0.340	0.086
	(0.337)	(0.322)	(0.625)	(0.322)
Observations	16,404	19,284	22,540	26,428
Log-likelihood	-93960	-22077	-82743	-83144
Wald chi2	5194	3971	1959	8708

Cluster robust standard errors in parentheses,
constants are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

(a) Selected Low-tech Input Industries

VARIABLES	(1)	(2)	(3)	(4)
l_lnI_i	0.290	0.350*	0.716	0.537
	(0.695)	(0.199)	(1.179)	(0.370)
l_lnI_o	0.051	0.087	0.549*	0.057
	(0.076)	(0.053)	(0.312)	(0.130)
l_lnE_high_i	-0.009	0.110***	0.062**	0.006
	(0.011)	(0.033)	(0.029)	(0.012)
l_lnE_high_ratio_i	-0.394	4.662***	1.650	-0.051
	(0.472)	(1.287)	(1.131)	(0.215)
l_lnE_high_o	0.012	-0.047***	-0.002	0.070***
	(0.014)	(0.012)	(0.018)	(0.022)
l_lnE_high_ratio_o	0.180	-1.940***	-0.560	2.671***
	(0.502)	(0.408)	(0.795)	(0.782)
rdint_i	-0.055***	0.007	0.031***	0.008
	(0.015)	(0.029)	(0.010)	(0.013)
rdint_o	-0.006	-0.026***	0.024*	-0.040***
	(0.009)	(0.007)	(0.014)	(0.013)
invint_i	-0.012*	0.014	-0.004	-0.006
	(0.007)	(0.013)	(0.017)	(0.007)
invint_o	0.005*	0.003	-0.016**	0.005
	(0.003)	(0.005)	(0.008)	(0.007)
lnempn_i	0.954	1.316***	-0.967	-1.343
	(0.993)	(0.287)	(1.917)	(0.900)
lnempn_o	0.356	0.027	-0.108	0.388
	(0.242)	(0.249)	(0.381)	(0.363)
researchers_i	-2.012	2.757***	2.353	3.996***
	(1.726)	(0.531)	(3.087)	(1.300)
researchers_o	2.402***	3.216***	1.398	2.341*
	(0.792)	(0.567)	(1.219)	(1.352)
lngdppc_i	1.207	0.445	21.133	-3.076***
	(0.871)	(0.536)	(13.478)	(0.902)
lngdppc_o	0.280	-1.042***	6.743**	-0.673
	(0.828)	(0.378)	(3.335)	(0.515)
lngdp_i	-0.632	-0.955***	-20.578	1.592***
	(0.457)	(0.266)	(13.077)	(0.595)
lngdp_o	-0.310	0.458*	-5.939*	0.414
	(0.513)	(0.234)	(3.431)	(0.365)
techdist_s	0.905	-3.332***	-0.758	-0.950
	(1.329)	(1.229)	(1.793)	(1.545)
techdist_cs	-2.212***	-2.390***	-0.472	-0.968
	(0.678)	(0.617)	(0.805)	(0.796)
techdist_c	-2.779***	-2.017**	-5.460***	-2.085*
	(0.811)	(0.885)	(1.977)	(1.094)
lndist	0.361	0.184	13.163	-0.031
	(0.252)	(0.132)	(8.149)	(0.184)
contig	1.005*	0.665**	9.806*	0.857**
	(0.547)	(0.324)	(5.818)	(0.423)
comlang_off	0.352	0.998***	12.492	0.539***
	(0.293)	(0.151)	(8.476)	(0.187)
colony	0.826	0.357	-4.199	-0.261
	(0.675)	(0.267)	(5.520)	(0.236)
c_pair	3.441***	3.149***	28.141**	2.612***
	(0.716)	(0.290)	(13.872)	(0.430)
s_pair	2.135***	-1.281**	0.249	0.769
	(0.674)	(0.553)	(1.821)	(1.336)
cs_pair	0.046	-0.122	0.359	0.157
	(0.478)	(0.346)	(1.143)	(0.482)
Observations	26,100	29,273	27,675	21,577
Log-likelihood	-350838	-173438	-515170	-143399
Wald chi2	10421	9507	7665	6620

Cluster robust standard errors in parentheses,
constants are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

(b) High-tech Input Industries

Table 4.A.8: Poisson Models for Different Input Industries with External High-tech and Low-tech R&D Stocks, Dependent Variable: Number of Forward Citations that Input Sector-country Receives from Output Sector-country

The column numbers refer to the following input industries: (a)(1) Agriculture, Hunting, Forestry and Fishing (2) Wood and Products of Wood and Cork (3) Pulp, Paper, Printing and Publishing (4) Construction; (b)(1) Chemicals and Chemical Products (2) Machinery, nec (3) Electrical and Optical Equipment (4) Transport Equipment.

VARIABLES	(1)	(2)	(3)	(4)
l_lnI_i	-0.027	-0.091***	0.319***	-0.116***
	(0.044)	(0.031)	(0.114)	(0.042)
l_lnI_o	0.088*	0.127***	0.289***	0.076*
	(0.049)	(0.039)	(0.096)	(0.042)
l_lnE_high_top_i	0.065	-0.010	0.781***	0.007*
	(0.268)	(0.199)	(0.207)	(0.004)
l_lnE_high_top_ratio_i	2.864**	0.015	1.244*	0.216
	(1.241)	(0.319)	(0.728)	(0.227)
l_lnE_high_top_o	0.012*	0.025***	0.022***	0.005
	(0.006)	(0.004)	(0.007)	(0.004)
l_lnE_high_top_ratio_o	-0.140	0.083	0.041	0.141
	(0.179)	(0.099)	(0.303)	(0.089)
rdint_i	0.233*	0.148	-0.035	-0.164
	(0.125)	(0.092)	(0.089)	(0.119)
rdint_o	-0.014	-0.013	0.005	-0.015
	(0.011)	(0.012)	(0.013)	(0.011)
invint_i	-0.008	0.002	-0.022	0.023
	(0.013)	(0.006)	(0.014)	(0.026)
invint_o	0.006	0.004	0.008	0.006
	(0.005)	(0.005)	(0.006)	(0.004)
lnempn_i	-0.324	0.061	2.615***	-0.451**
	(0.312)	(0.103)	(0.363)	(0.223)
lnempn_o	0.495***	0.321***	0.743***	0.174
	(0.157)	(0.079)	(0.246)	(0.180)
researchers_i	-2.139*	3.832***	3.512***	4.400***
	(1.281)	(0.381)	(1.159)	(0.641)
researchers_o	0.612	2.393***	2.337**	2.355***
	(0.838)	(0.424)	(1.035)	(0.687)
lngdppc_i	0.557	-1.236***	-0.046	-2.138***
	(0.993)	(0.300)	(0.493)	(0.332)
lngdppc_o	0.465	-0.978***	0.817	-0.967***
	(0.686)	(0.245)	(0.520)	(0.355)
lngdp_i	0.752***	0.888***	-2.137***	1.505***
	(0.277)	(0.089)	(0.426)	(0.227)
lngdp_o	0.219	0.312***	-0.217	0.601***
	(0.156)	(0.092)	(0.308)	(0.187)
techdist_s	-2.051**	-4.395***	-5.239***	-3.530***
	(0.909)	(0.536)	(1.390)	(0.946)
techdist_cs	-0.085	-0.190	0.518	-0.408
	(0.960)	(0.424)	(0.831)	(0.432)
techdist_c	-0.807	-5.374***	-2.380**	-2.796***
	(1.105)	(0.564)	(1.144)	(0.488)
lndist	0.040	-0.438***	-0.263***	-0.144***
	(0.151)	(0.041)	(0.099)	(0.045)
contig	0.850***	-0.334***	0.265	0.599***
	(0.250)	(0.110)	(0.204)	(0.120)
comlang_off	0.360*	0.764***	0.571**	0.740***
	(0.215)	(0.104)	(0.248)	(0.087)
colony	-0.754**	-0.272**	-0.797***	-0.263**
	(0.383)	(0.106)	(0.206)	(0.117)
c_pair	3.148***	1.125***	2.561***	2.422***
	(0.436)	(0.144)	(0.380)	(0.175)
s_pair	-0.088	-1.810***	-1.575*	-0.992*
	(0.788)	(0.472)	(0.915)	(0.597)
cs_pair	0.223	0.342	0.333	0.080
	(0.337)	(0.342)	(0.653)	(0.313)
Observations	16,404	19,284	22,540	26,428
Log-likelihood	-93733	-22139	-83849	-83224
Wald chi2	5507	3801	1828	9392

Cluster robust standard errors in parentheses,
constants are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

(a) Selected Low-tech Input Industries

VARIABLES	(1)	(2)	(3)	(4)
l_lnI_i	0.255	0.391**	0.233	0.450
	(0.692)	(0.184)	(1.071)	(0.354)
l_lnI_o	0.051	0.096*	0.613*	0.088
	(0.073)	(0.058)	(0.371)	(0.108)
l_lnE_high_top_i	-0.002	-0.005	0.017*	0.009
	(0.006)	(0.007)	(0.010)	(0.010)
l_lnE_high_top_ratio_i	0.014	0.213	-0.072	0.179
	(0.015)	(0.435)	(0.155)	(0.127)
l_lnE_high_top_o	0.009	-0.008*	0.024**	0.001
	(0.008)	(0.005)	(0.010)	(0.009)
l_lnE_high_top_ratio_o	0.012	-0.477***	-0.359	0.106*
	(0.017)	(0.138)	(0.230)	(0.065)
rdint_i	-0.055***	-0.000	0.033***	0.004
	(0.015)	(0.034)	(0.009)	(0.013)
rdint_o	-0.006	-0.030***	0.022*	-0.035***
	(0.008)	(0.009)	(0.012)	(0.012)
invint_i	-0.012*	0.013	-0.009	-0.006
	(0.007)	(0.013)	(0.016)	(0.007)
invint_o	0.005**	0.004	-0.012	0.005
	(0.003)	(0.005)	(0.007)	(0.006)
lnempn_i	0.994	1.371***	0.273	-1.192
	(0.994)	(0.290)	(1.454)	(0.884)
lnempn_o	0.326	0.073	-0.257	0.274
	(0.235)	(0.263)	(0.453)	(0.325)
researchers_i	-2.078	2.713***	3.988*	4.144***
	(1.754)	(0.548)	(2.418)	(1.302)
researchers_o	2.412***	3.248***	2.091*	2.088*
	(0.797)	(0.564)	(1.177)	(1.254)
lngdppc_i	1.307	0.701	6.631	-3.066***
	(0.865)	(0.558)	(9.494)	(0.918)
lngdppc_o	0.311	-1.038***	1.401	-0.645
	(0.836)	(0.403)	(2.273)	(0.518)
lngdp_i	-0.639	-1.080***	-7.300	1.619***
	(0.459)	(0.297)	(9.478)	(0.598)
lngdp_o	-0.315	0.378	-1.024	0.522
	(0.534)	(0.279)	(1.841)	(0.334)
techdist_s	0.868	-3.519***	-0.291	-1.812
	(1.344)	(1.198)	(1.768)	(1.554)
techdist_cs	-2.227***	-2.430***	-0.905	-1.030
	(0.684)	(0.683)	(0.691)	(0.924)
techdist_c	-2.803***	-2.277**	-4.656***	-2.224**
	(0.822)	(1.044)	(1.502)	(1.052)
lndist	0.373	0.237	5.759	-0.068
	(0.266)	(0.168)	(6.611)	(0.174)
contig	1.018*	0.715*	9.040	0.771**
	(0.565)	(0.398)	(8.407)	(0.362)
comlang_off	0.329	1.031***	2.332	0.564***
	(0.305)	(0.173)	(4.385)	(0.162)
colony	0.846	0.378	-1.305	-0.299
	(0.690)	(0.313)	(4.952)	(0.211)
c_pair	3.461***	3.245***	12.560*	2.522***
	(0.770)	(0.368)	(7.453)	(0.392)
s_pair	2.160***	-1.476***	-3.676	-0.080
	(0.661)	(0.549)	(3.461)	(1.225)
cs_pair	0.060	-0.137	-1.873	0.182
	(0.490)	(0.361)	(2.543)	(0.445)
Observations	26,100	29,273	27,675	21,577
Log-likelihood	-350902	-174491	-501945	-144203
Wald chi2	10118	9610	5613	6646

Cluster robust standard errors in parentheses,
constants are suppressed.
*** p<0.01, ** p<0.05, * p<0.1

(b) High-tech Input Industries

Table 4.A.9: Poisson Models for Different Input Industries with External 'Top' High-tech and Low-tech R&D stocks, Dependent Variable: Number of Forward Citations that Input Sector-country Receives from Output Sector-country

The column numbers refer to the following input industries: (a)(1) Agriculture, Hunting, Forestry and Fishing (2) Wood and Products of Wood and Cork (3) Pulp, Paper, Printing and Publishing (4) Construction; (b)(1) Chemicals and Chemical Products (2) Machinery, nec (3) Electrical and Optical Equipment (4) Transport Equipment.

4.B Appendix

4.B.1 Derivation of the Empirical Model

Let

$$Q_{ci,si,t} = I^{\alpha_1 i}_{ci,si,t-1} E^{\alpha_2 i}_{ci,si,t-1} e^{v_{ci}} e^{v_{si}} e^{v_t} \tag{4.B.1}$$

a log-linear production function where $Q_{ci,si,t}$ is technological output ('innovations') in country ci, industry si in year t. $I_{ci,si,t}$ is the knowledge stock that is accumulated 'own' knowledge that accrues to the focal sector-country, $E_{ci,si,t}$ is potentially accessible knowledge that is accumulated in other sector-countries (external knowledge), $e^{v_{ci}}$, $e^{v_{si}}$, and e^{v_t} capture country, industry and time specific effects. We assume that some fraction of $Q_{ci,si,t}$ is patented $(1/\lambda_{ci,si,t})$ and that the number of new patents is given by

$$P_{ci,si,t} = 1/\lambda_{ci,si,t} Q_{ci,si,t} = 1/\lambda_{ci,si,t} I^{\alpha_1}_{ci,si,t-1} E^{\alpha_2}_{ci,si,t-1} e^{v_{ci}} e^{v_{si}} e^{v_t} \tag{4.B.2}$$

A fraction of the patents in ci, si is cited by co, so and $C_{ci,si,co,so,t}$ is the number of citations that the input sector-country ci, si receives from co, so. Following Peri (2005), we assume that citations are a noisy indicator of actual outflows $\Phi_{ci,si,co,so,t}$ and include an error term:

$$C_{ci,si,co,so,t} = \Phi_{ci,si,co,so,t} e^{\epsilon_{ci,si,co,so,t}} \tag{4.B.3}$$

The relative intensity of knowledge flows that goes from ci, si to co, so relative to the number of innovations in ci, si times the number of innovations in co, so is given by

$$\theta_{ci,si,co,so,t} = \frac{\Phi_{ci,si,co,so,t}}{Q_{ci,si,t} Q_{co,so,t}} = \frac{C_{ci,si,co,so,t}}{\lambda_{ci,si,t} P_{ci,si,t} \lambda_{co,so,t} P_{co,so,t} e^{\epsilon_{ci,si,co,so,t}}} \tag{4.B.4}$$

if we re-arrange both the first equality of (4.B.2) and (4.B.3) and insert them afterwards. Again following Peri, the relative intensity of knowledge flows between 'sending' sector-countries and 'receiving' sector-countries can be also written as an exponential function dependent on a host of geographic and technological variables:

$$\theta_{ci,si,co,so,t} = exp(\boldsymbol{x'_{ci,si,co,so,t}\beta}) \tag{4.B.5}$$

Combining (4.B.4) and (4.B.5) and after re-arranging, $C_{ci,si,co,so,t}$ can be written as exponential function of geographic and technological variables $\boldsymbol{x'_{ci,si,co,so,t}}$ and patents in ci, si and co, so:

$$C_{ci,si,co,so,t} = exp(\boldsymbol{x'_{ci,si,co,so,t}}\boldsymbol{\beta} + lnP_{ci,si,t} + lnP_{co,so,t} + ln\lambda_{ci,si,t} + ln\lambda_{co,so,t} + \epsilon_{ci,si,co,so,t}) \quad (4.B.6)$$

After inserting (4.B.2), we receive a function that depends on input and output knowledge stocks, external spillovers, input and output specific effects, time effects, an error term and additional variables of interest $\boldsymbol{x'_{ci,si,co,so,t}}$:

$$\begin{aligned} C_{ci,si,co,so,t} = exp(\boldsymbol{x'_{ci,si,co,so,t}}\boldsymbol{\beta} + \alpha_{1i}lnI_{ci,si,t-1} + \alpha_{1o}lnI_{co,so,t-1} + \\ \alpha_{2i}lnE_{ci,si,t-1} + \alpha_{2o}lnE_{co,so,t-1} + v_{ci} + v_{co} + v_{si} + v_{so} + v_t + \xi_{ci,si,co,so,t}) \end{aligned} \quad (4.B.7)$$

4.B.2 Further Tables & Figures

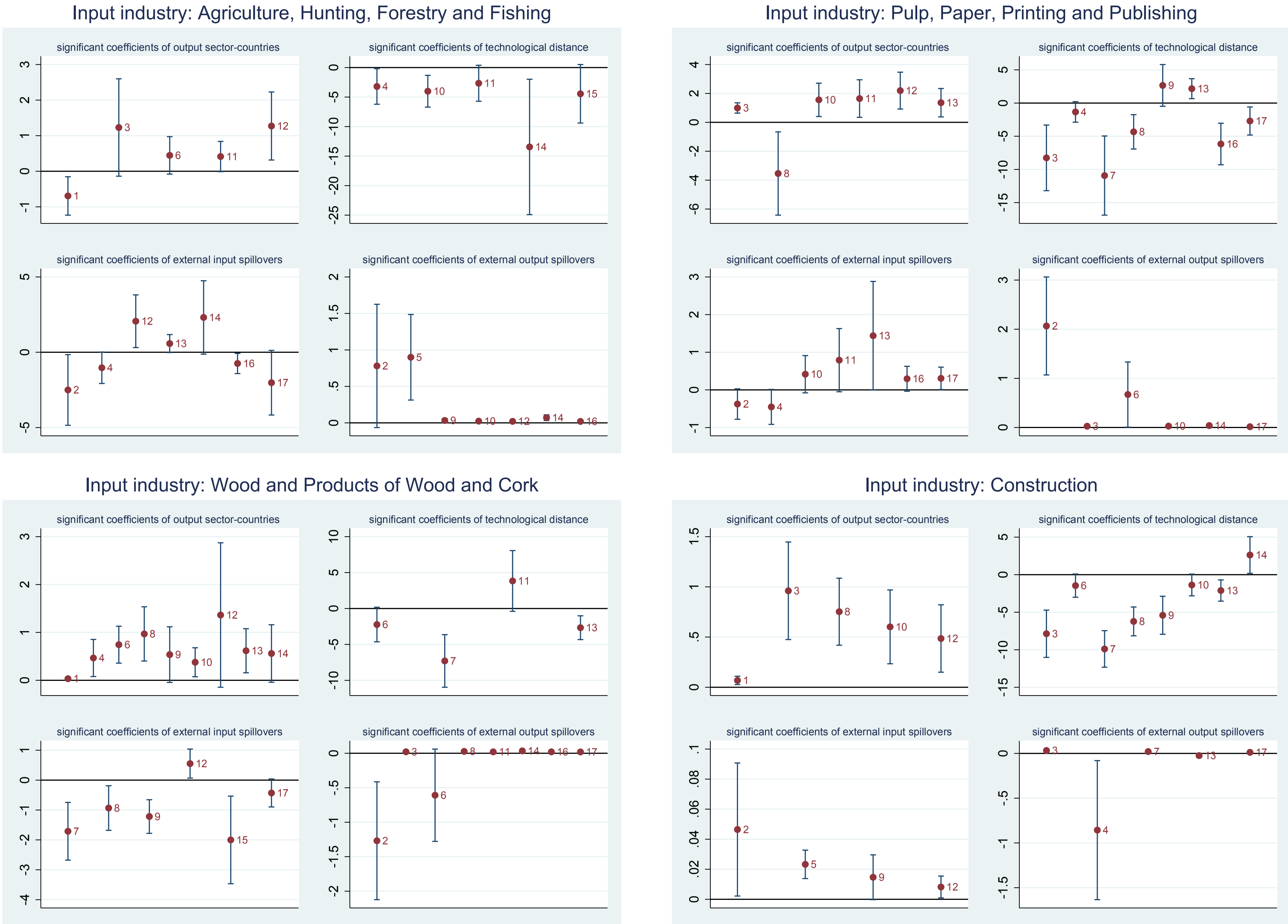

Figure 4.B.1: Significant Coefficients and 95% Confidence Intervals for Selected Low-tech Input Industries

The numbers in the diagrams refer to the following output industries: (1) Agriculture, Hunting, Forestry and Fishing (2) Mining and Quarrying (3) Food, Beverages and Tobacco (4) Textiles and Leather Products (5) Wood and Products of Wood and Cork (6) Pulp, Paper, Printing and Publishing (7) Coke, Refined Petroleum and Nuclear Fuel (8) Chemicals and Chemical Products (9) Rubber and Plastics (10) Other Non-Metallic Mineral (11) Basic Metals and Fabricated Metal (12) Machinery, nec (13) Electrical and Optical Equipment (14) Transport Equipment (15) Manufacturing, Nec; Recycling (16) Electricity, Gas and Water Supply (17) Construction.

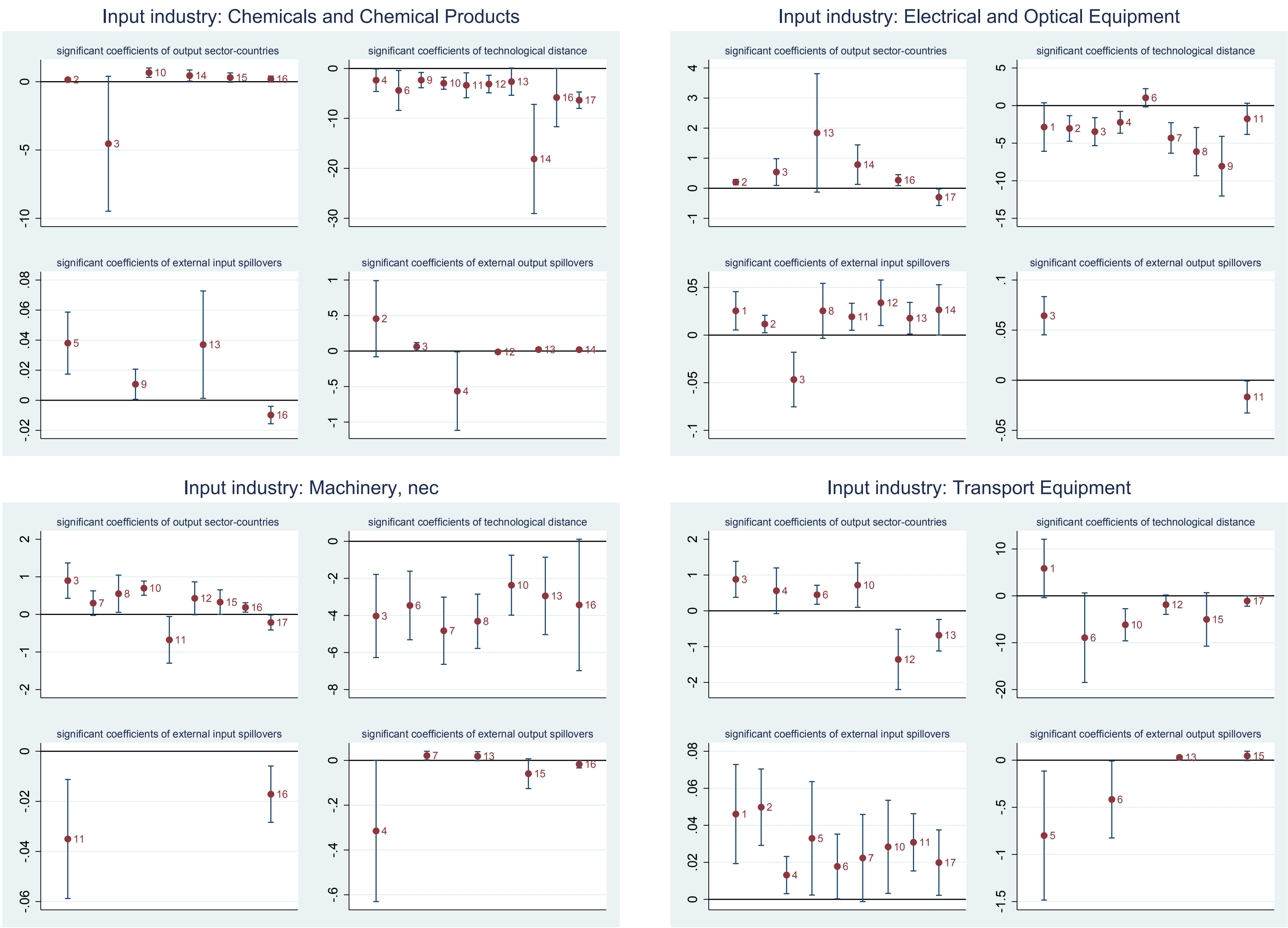

Figure 4.B.2: Significant Coefficients and 95% Confidence Intervals for High-tech Input Industries

The numbers in the diagrams refer to the following output industries: (1) Agriculture, Hunting, Forestry and Fishing (2) Mining and Quarrying (3) Food, Beverages and Tobacco (4) Textiles and Leather Products (5) Wood and Products of Wood and Cork (6) Pulp, Paper, Printing and Publishing (7) Coke, Refined Petroleum and Nuclear Fuel (8) Chemicals and Chemical Products (9) Rubber and Plastics (10) Other Non-Metallic Mineral (11) Basic Metals and Fabricated Metal (12) Machinery, nec (13) Electrical and Optical Equipment (14) Transport Equipment (15) Manufacturing, Nec; Recycling (16) Electricity, Gas and Water Supply (17) Construction.

Table 4.B.1: Logit Models, Dependent Variable: Incidence of Forward Citations between Input Sector-country and Output Sector-country

VARIABLES	(1) Logit fw_citation_d	(2) Logit fw_citation_d	(3) Logit fw_citation_d
l_lnI_i	0.004*	0.004*	0.004*
	(0.002)	(0.002)	(0.002)
l_lnI_o	0.013***	0.013***	0.013***
	(0.002)	(0.002)	(0.002)
l_lnE_i	-0.006***		
	(0.002)		
l_lnE_o	0.010***		
	(0.002)		
l_lnE_high_i		-0.004**	
		(0.002)	
l_lnE_high_ratio_i		0.068***	
		(0.024)	
l_lnE_high_o		0.013***	
		(0.002)	
l_lnE_high_ratio_o		0.097***	
		(0.021)	
l_lnE_high_top_i			-0.006***
			(0.002)
l_lnE_high_top_ratio_i			-0.010**
			(0.005)
l_lnE_high_top_o			0.011***
			(0.002)
l_lnE_high_top_ratio_o			0.039***
			(0.005)
rdint_i	0.017***	0.017***	0.017***
	(0.001)	(0.001)	(0.001)
rdint_o	0.013***	0.013***	0.013***
	(0.001)	(0.001)	(0.001)
invint_i	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
invint_o	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
lnempn_i	0.043***	0.043***	0.042***
	(0.011)	(0.011)	(0.011)
lnempn_o	0.011	0.011	0.012
	(0.011)	(0.011)	(0.011)
researchers_i	1.182***	1.179***	1.182***
	(0.142)	(0.142)	(0.142)
researchers_o	0.339**	0.335**	0.325**
	(0.149)	(0.149)	(0.149)
lngdppc_i	-4.884***	-4.862***	-4.894***
	(0.267)	(0.267)	(0.267)
lngdppc_o	-2.163***	-2.134***	-2.111***
	(0.268)	(0.268)	(0.268)
lngdp_i	3.257***	3.236***	3.263***
	(0.250)	(0.250)	(0.250)
lngdp_o	2.095***	2.067***	2.053***
	(0.254)	(0.254)	(0.254)
techdist_s	-1.342***	-1.344***	-1.372***
	(0.044)	(0.044)	(0.044)
techdist_cs	-0.334***	-0.331***	-0.320***
	(0.039)	(0.039)	(0.039)
techdist_c	-1.112***	-1.115***	-1.116***
	(0.062)	(0.062)	(0.062)
lndist	-0.064***	-0.064***	-0.064***
	(0.017)	(0.017)	(0.017)
contig	0.311***	0.311***	0.311***
	(0.026)	(0.026)	(0.026)
comlang_off	0.365***	0.365***	0.365***
	(0.026)	(0.026)	(0.026)
colony	-0.377***	-0.377***	-0.377***
	(0.030)	(0.030)	(0.030)
c_pair	2.981***	2.981***	2.982***
	(0.062)	(0.062)	(0.062)
s_pair	-0.387***	-0.385***	-0.396***
	(0.032)	(0.032)	(0.032)
cs_pair	-0.496**	-0.495**	-0.493**
	(0.210)	(0.210)	(0.210)
Observations	787,260	787,260	787,260
Log-likelihood	-326639	-326627	-326588
Wald chi2	132218	132285	132304

Cluster robust standard errors in parentheses,
the constant is suppressed
*** p<0.01, ** p<0.05, * p<0.1

5 On Patent Citations and its Fundamentals*

5.1 Introduction

Patents and their citations are particularly relevant for two reasons. First, patents reflect technological change as a key driver of economic growth (see, e.g., Jaffe and Trajtenberg 2002). Second, patent citations reflect the quality and value of innovations (see, e.g., Hall et al. 2005; Lanjouw and Schankerman 2004; Trajtenberg 1990). Understanding how they are linked to complementary factors of an economy is of key importance to policy makers.

Patent-activity distributions can be relatively well described by power laws. For example, O'Neale and Hendy (2012) document this for patent filings across countries, Silverberg and Verspagen (2007) provide evidence for patent citations, and Pakes (1986); Scherer (2000); Scherer and Harhoff (2000) illustrate it for patent values. The power-law form of patent citations and values suggests that only a small fraction of patented inventions yields high technical and economic value. We do not fully understand yet economic fundamentals behind citation power laws as such. The saliency of certain features of patent-citation distributions suggests that interesting latent processes might generate them (see Scherer 2000), but it is not yet well understood how fundamentals such as R&D expenditures, market size, or openness contribute to the concentration or dispersion of patent citations across countries and sectors.

This paper characterizes the citations for all 3.7 mn. patent-family applications over the period 1995-2005 across 17 industries and 34 countries by power-law regressions and then determines to which extent the power-law characteristics relate to economic and institutional characteristics. This is achieved by combining estimates from power-law regressions with a subsequent

*This chapter is co-authored with Peter Egger & Martin Wörter. The Chapter was published in Economics Letters (https://doi.org/10.1016/j.econlet.2016.08.011). Elsevier granted the right to reuse it in a thesis/dissertation both in printed and electronic format.

fixed-effects quantiles regression analysis on the power-law coefficients. Key findings are that a greater R&D intensity and outward foreign direct investment in a country and sector suggest a greater dispersion of patent citations and, hence, patent values.

5.2 Estimating Power-law Regressions for Patent Citations

Let us denote the shape parameter of the Pareto distribution of patent citations in country i, sector s, and time t by k_{ist}. Moreover, let $1 - \Pr_{pist}$ denote the probability for a patented invention of having at least as many citations as patented inventions in the p-th percentile of the distribution, and let $rank_p$ denote the corresponding rank. Hence, $1 - \Pr_{99ist}$ is associated with $rank_1$. Let us refer to the average number of citations of all patented inventions up until the p-th percentile of the distribution in country i, sector s, and time t by $\bar{c}_{pist}$, which is computed cumulatively in the first percentile, in the first-plus-second percentile, etc. Then, k_{ist} can be estimated per $\{ist\}$ by a generalized-linear exponential-family model of the form

$$\bar{c}_{pist} = \exp(\underbrace{\alpha_{ist}}_{=-(1/k_{ist})} \ln rank_{pist} + \mu_{ist} + error_{pist}), \tag{5.2.1}$$

where μ_{ist} is a fixed country-sector-time effect, and $rank_{pist} \in \{1, ..., 99\}$. There are 99 degrees of freedom to estimate each one of the parameters k_{ist}. A parameter α_{ist} that is larger (smaller) in absolute value suggests that citations drop off more (less) severely as we move down the citation ranks. Hence, the concentration of citations rises with the absolute value of α_{ist}. We consider two implementations of (5.2.1), one where α_{ist} is estimated in a pooled fashion for all centiles p per $\{ist\}$, and one where we relax pooling to acknowledge a deviation from the single-parameter Pareto power law, estimating slope parameters per decile d among the centiles p, α^d_{ist}. The latter allows for a deviation from linearity in the otherwise linear index in (5.2.1). We estimate (5.2.1) based on the universe of all 19.2 mn. patent family citations reported in the European Patent Office (EPO) Patstat data for all 34 OECD countries[1] and 17 sectors[2] for

[1] The OECD comprises the following countries: Australia, Austria, Belgium, Canada, Chile, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

[2] Agriculture, Hunting, Forestry and Fishing; Mining and Quarrying; Food, Beverages and Tobacco; Textiles and Textile Products, Leather and Footwear; Wood and Products of Wood and Cork; Pulp, Paper, Printing and Publishing; Coke, Refined Petroleum and Nuclear Fuel; Chemicals and Chemical Products; Rubber and Plastics; Other Non-Metallic Mineral; Basic Metals and Fabricated Metal; Machinery; Electrical and Optical Equipment; Transport Equipment; Other Manufacturing and Recycling; Electricity, Gas and Water Supply; Construction.

patent-family applications between 1995-2005 per country, sector, and year that occur within 5 years after publication of the cited patent family application. The sector assignment is based on the concordance tables from Lybbert and Zolas (2014). The number of forward citations per patent and industry is weighted based on the technological fields (International Patent Classification on subclass level) a patent family belongs to, yielding 6,358 triplets $\{ist\}$. Industry and country data on possible determinants are taken from the OECD STAN database, the World Bank World Development Indicators, and the United Nations Conference on Trade and Development (UNCTADSTAT).

The upper four panels in Figure 5.A.1 summarize the relationship postulated in (5.2.1) for two exemplary pairs of countries (Germany and the Netherlands) and sectors (Electrical/Optical Equipment and Wood Products) for $\hat{\alpha}_{ist}$ in the average year in 1995-2005. The lower two panels portray all estimates $\hat{\alpha}_{ist}$ as well as their demeaned value, $\breve{\hat{\alpha}}_{ist}$ (subtracting the average per is over time) by way of kernel-density plots. We observe some deviation from the (log-linear) power law which varies across sectors, supporting an estimation of decile-specific α^d_{ist} for better approximation. Moreover, the distribution of $\hat{\alpha}_{ist}$ looks twin-peaked and asymmetric due to time-invariant, country-sector-specific factors.

5.3 Explaining Power-law Regression Coefficients in Fixed-effects Quantile Regressions

We are particularly interested in determining the concentration of patent citations by relating $\hat{\alpha}_{ist}$ and $\hat{\alpha}^d_{ist}$ to economic fundamentals. We treat them as estimated, heteroskedastic dependent variables (see Saxonhouse 1977), explaining them by a number of lagged time-variant fundamentals, X_{ist-1}, in linear (see Baltagi 2008) and quantiles fixed-effects regressions (see Canay 2011). These are three measures capturing the knowledge environment in a country (log patent stock per country and sector, the R&D intensity, the share of people with tertiary school enrolment per country and sector); three variables measuring sector size (log employment), tangible-investment intensity (investment over sales), and profitability (log operating profits from sales) per country and sector; and four measures of openness (export intensity and import intensity per country and sector, outward and inward FDI stocks relative to GDP per country). As we allow for some endogeneity of all determinants of $\hat{\alpha}_{ist}$, we include residuals from first-step regressions in X_{ist-1} which are obtained from regressions of fundamentals on lagged differences of those determinants (see the footnote to Table 5.A.3; notice that such a control function approach is consistent with the idea in Chernozhukov and Hansen (2008). The proposed models are

$$\hat{\beta} \quad = \quad \arg\min_{\beta,\breve{\lambda}} \mathbb{E}[(\breve{\hat{\alpha}}_{ist} - \breve{X}_{ist-1}\beta - \breve{\lambda}_{is})^2], \qquad \hat{\beta}_q = \arg\min_{\beta_q} \mathbb{E}[\rho_q(\breve{\hat{\alpha}}_{ist} - \breve{X}_{ist-1}\beta_q - \breve{\hat{\lambda}}_{is})], \tag{5.3.1}$$

where "$\breve{}$" denotes quantities that are adjusted for (power-law-regression-)imprecision, $\breve{\lambda}_{is}$ are is-specific fixed effects, q indicates a quantile, ρ_q is the *check* or *asymmetric absolute loss* function (see Canay 2011; Wooldridge 2010), and $\breve{\hat{\lambda}}_{is}$ are step-1 estimates of fixed effects which are not estimated in Step 2 of the quantiles model, where $\hat{\beta}_q$ is estimated. We conduct all estimations for both pooled $\breve{\hat{\alpha}}_{ist}$ and decile-specific $\breve{\hat{\alpha}}^d_{ist}$. A negative (positive) parameter on one of these variables suggests that a higher value of that variable is associated with a bigger (smaller) concentration of citations in a sector, country, and subsequent year.

Tables 5.A.2 and 5.A.3 summarize the results and are organized in four blocs. These blocs refer to all estimates of $\breve{\hat{\alpha}}_{ist}$ (upper left) and, using super-script d to denote deciles of $\breve{\hat{\alpha}}_{ist}$, the subsamples of first-decile (upper right), fifth-decile (lower left), and ninth-decile estimates (lower right) of $\breve{\hat{\alpha}}^d_{ist}$, respectively. Each of the four blocs in Table 5.A.2 and 5.A.3 contains six columns for the linear model and the 10th, 25th, 50th, 75th, and 90th percentiles for the quantiles models. As already mentioned, in Table 5.A.3, we apply a Control Function Approach to address potential endogeneity (see Terza et al. 2008; Wooldridge 2015). In the following, we only comment on results from the Control Function Approach.[3]

The results suggest that the association between the considered fundamentals and the power-law of patent citations should rely on nonlinear estimation techniques such as quantiles models rather than linear models: some of the parameter estimates differ between these types of models even in a qualitative dimension, and the quantitative differences are often large relative to the quantiles point estimates. Specifically, bigger patent stocks tend to be associated with a more dispersed pattern of citations (and values) of patents, but this effect is estimated to be weaker by the quantiles model than the linear one (see, e.g., the top-left panel in Table 5.A.3). A higher R&D intensity of a country and sector also tends to reduce the concentration of patent citations (and values) on average and for two quantiles of the parameter distribution in the first

[3]Tables 5.A.4 and 5.A.5 provide the same set of results for the pooled parameter estimated with a Generalized Linear Model instead of OLS as a robustness test.

decile of the distribution of citations (see the two top blocks of Table 5.A.3).[4] The effects of tertiary school enrolment are less clear-cut. These findings suggest that policy measures which are aimed at stimulating R&D and patenting appear to reduce the concentration of patent values across inventors and owners. A higher profitability of a sector and country tends to raise the concentration of patent citations on average and in the first decile of the distribution of citations (see the parameters in the top blocks of Table 5.A.3).

Among the openness measures, both outward and inward FDI appear to have a coherent association with the citation power law: outward FDI tends to reduce the concentration of patent citations (and values) in a country throughout the distribution, whereas inward FDI tends to increase the concentration. One reason for this result might be the relocation of highly-cited-and-valued patents by multinational firms to tax havens.

[4] A higher R&D intensity leads to more independent trials which increases the probability that several valuable patents are highly cited across sectors and countries.

Appendix

5.A Tables & Figures

Table 5.A.1: Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Log patent stock$_{ist-1}$	1,804	7.012	1.469	4.148	11.441
R&D/sales$_{ist-1}$	1,804	1.065	1.844	0.000	12.790
Share tertiary education$_{it-1}$	1,804	22.309	7.419	8.900	44.000
Log no. employees$_{ist-1}$	1,804	11.104	1.347	7.454	14.815
Investments/sales$_{ist-1}$	1,804	7.555	5.591	0.668	57.203
Log net oper. profit from sales$_{ist-1}$	1,804	21.424	2.110	13.916	30.722
Export/sales$_{ist-1}$	1,804	51.369	88.270	0.000	1694.414
Import/sales$_{ist-1}$	1,804	75.385	194.988	0.000	3076.515
Outward FDI stock/GDP$_{it-1}$	1,804	29.219	25.696	0.431	103.218
Inward FDI stock/GDP$_{it-1}$	1,804	36.092	30.012	2.662	148.671

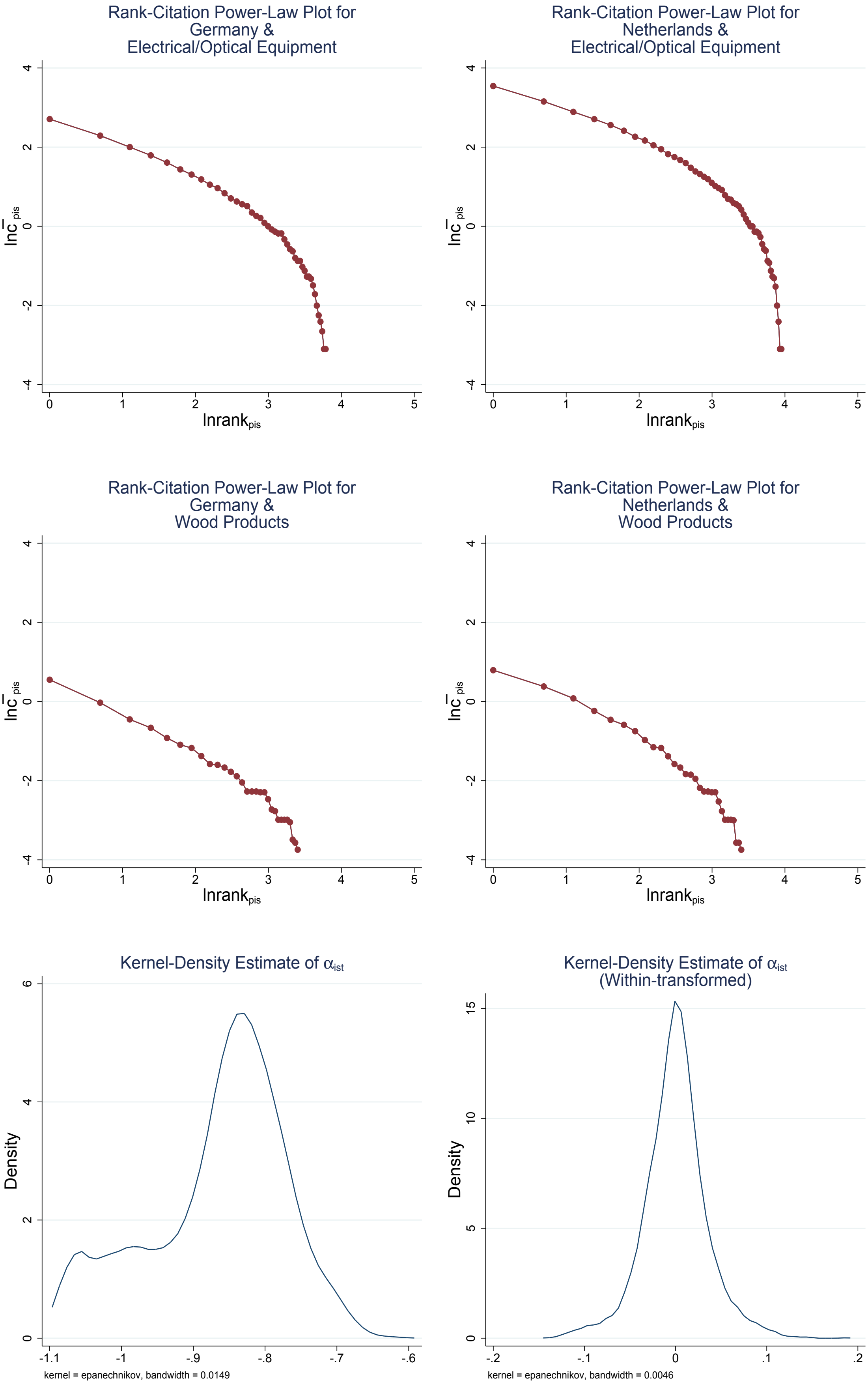

Figure 5.A.1: Rank-Citation Power-Law Plots and Kernel-Density Estimates

Table 5.A.2: Estimation Results

	Results for Pooled Parameter $\check{\alpha}_{ist}$						Results for Pooled First-decile Parameter $\check{\alpha}^1_{ist}$					
		Quantiles Regression Results for Percentile						Quantiles Regression Results for Percentile				
Covariate	Linear	10	25	50	75	90	Linear	10	25	50	75	90
Log patent stock$_{ist-1}$	0.013238	0.023176 ***	0.023896 ***	0.023051 ***	0.021009 ***	0.018507 ***	0.020981 ***	0.008516 ***	0.008985 ***	0.009029 ***	0.008008 ***	0.006606 ***
	(0.010348)	(0.002103)	(0.001603)	(0.001686)	(0.002005)	(0.001969)	(0.005486)	(0.000808)	(0.000746)	(0.000756)	(0.000808)	(0.000971)
R&D/sales$_{ist-1}$	-0.016731 *	0.001602 *	0.001458 *	0.001750 **	0.001545 **	0.001257 *	-0.002711	0.000287	0.000274	0.000678 *	0.000721 *	0.000786
	(0.009360)	(0.000867)	(0.000803)	(0.000686)	(0.000778)	(0.000734)	(0.005442)	(0.000399)	(0.000288)	(0.000356)	(0.000387)	(0.000577)
Share tertiary education$_{it-1}$	-0.002775 ***	0.000727 *	0.000957 ***	0.001059 ***	0.001247 ***	0.001457 ***	-0.004290 **	0.000178	0.000227 *	0.000378 ***	0.000548 ***	0.000675 ***
	(0.000926)	(0.000372)	(0.000319)	(0.000362)	(0.000349)	(0.000351)	(0.001677)	(0.000136)	(0.000116)	(0.000125)	(0.000144)	(0.000185)
Log no. employees$_{ist-1}$	-0.078615 ***	-0.011396 ***	-0.014610 ***	-0.016175 ***	-0.016464 ***	-0.017123 ***	-0.075065 ***	-0.004212 ***	-0.005036 ***	-0.004944 ***	-0.004574 ***	-0.004406 ***
	(0.008241)	(0.003018)	(0.002184)	(0.002423)	(0.002616)	(0.002582)	(0.006346)	(0.001302)	(0.000923)	(0.000981)	(0.000979)	(0.001032)
Investments/sales$_{ist-1}$	-0.000625	-0.000639	-0.000657	-0.000480	-0.000543	-0.000854 **	-0.001298 *	-0.000325	-0.000248	-0.000180	-0.000228	-0.000243
	(0.000689)	(0.000543)	(0.000490)	(0.000415)	(0.000358)	(0.000359)	(0.000754)	(0.000240)	(0.000207)	(0.000164)	(0.000173)	(0.000167)
Log net oper. profit from sales$_{ist-1}$	-0.008440	-0.005051 ***	-0.004456 ***	-0.004163 ***	-0.003712 ***	-0.002528 **	-0.009026 ***	-0.001075 *	-0.000672 *	-0.000732 *	-0.000806 *	-0.000624
	(0.005411)	(0.001264)	(0.001040)	(0.001099)	(0.001125)	(0.001127)	(0.003138)	(0.000580)	(0.000398)	(0.000397)	(0.000475)	(0.000537)
Export/sales$_{ist-1}$	0.000056	0.000113 **	0.000081 *	0.000042	0.000004	-0.000011	-0.000273	0.000045 *	0.000033 ***	0.000020	-0.000004	-0.000022 *
	(0.000283)	(0.000054)	(0.000047)	(0.000040)	(0.000034)	(0.000032)	(0.000201)	(0.000027)	(0.000012)	(0.000013)	(0.000012)	(0.000013)
Import/sales$_{ist-1}$	-0.000311 ***	-0.000088 ***	-0.000074 ***	-0.000059 ***	-0.000042 **	-0.000040 ***	-0.000283 ***	-0.000032 **	-0.000025 ***	-0.000021 ***	-0.000012 **	-0.000006
	(0.000112)	(0.000031)	(0.000021)	(0.000020)	(0.000017)	(0.000015)	(0.000088)	(0.000016)	(0.000007)	(0.000006)	(0.000005)	(0.000006)
Outward FDI stock/GDP$_{it-1}$	-0.000164	0.001198 ***	0.001113 ***	0.001069 ***	0.000970 ***	0.000913 ***	0.000977 *	0.000357 ***	0.000337 ***	0.000322 ***	0.000369 ***	0.000496 ***
	(0.000335)	(0.000102)	(0.000108)	(0.000090)	(0.000095)	(0.000113)	(0.000575)	(0.000050)	(0.000043)	(0.000043)	(0.000054)	(0.000070)
Inward FDI stock/GDP$_{it-1}$	0.000616 **	-0.000457 ***	-0.000387 ***	-0.000292 ***	-0.000194 ***	-0.000162 *	0.000108	-0.000140 ***	-0.000101 ***	-0.000071 **	-0.000060	-0.000053
	(0.000300)	(0.000086)	(0.000097)	(0.000069)	(0.000069)	(0.000097)	(0.000256)	(0.000040)	(0.000036)	(0.000035)	(0.000042)	(0.000055)
No. of countries i	20	20	20	20	20	20	20	20	20	20	20	20
No. of sectors s	16	16	16	16	16	16	16	16	16	16	16	16
No. of years t	10	10	10	10	10	10	10	10	10	10	10	10

	Results for Pooled Fifth-decile Parameter $\check{\alpha}^5_{ist}$						Results for Pooled Ninth-decile Parameter $\check{\alpha}^9_{ist}$					
		Quantiles Regression Results for Percentile						Quantiles Regression Results for Percentile				
Covariate	Linear	10	25	50	75	90	Linear	10	25	50	75	90
Log patent stock$_{ist-1}$	0.011459	0.007123 ***	0.007830 ***	0.008376 ***	0.008832 ***	0.006787 ***	0.006009	0.007155 ***	0.006964 ***	0.006154 ***	0.004839 ***	0.003554 ***
	(0.020508)	(0.000727)	(0.000632)	(0.000592)	(0.000709)	(0.000659)	(0.012670)	(0.000732)	(0.000701)	(0.000572)	(0.000554)	(0.000575)
R&D/sales$_{ist-1}$	-0.024565	0.000217	-0.000018	-0.000188	-0.000288	-0.000340	-0.002750	0.000442	0.000534 *	0.000343	0.000349	0.000476
	(0.015443)	(0.000297)	(0.000260)	(0.000279)	(0.000372)	(0.000424)	(0.009083)	(0.000332)	(0.000293)	(0.000336)	(0.000413)	(0.000422)
Share tertiary education$_{it-1}$	-0.003636	-0.000067	0.000012	0.000165	0.000372 ***	0.000522 ***	-0.006513 ***	-0.000019	-0.000009	0.000052	-0.000080	-0.000233 *
	(0.001659)	(0.000087)	(0.000109)	(0.000109)	(0.000102)	(0.000147)	(0.001250)	(0.000132)	(0.000110)	(0.000139)	(0.000113)	(0.000123)
Log no. employees$_{ist-1}$	-0.090841 ***	-0.003439 ***	-0.003800 ***	-0.004321 ***	-0.004823 ***	-0.004825 ***	-0.067212 ***	-0.005257 ***	-0.004638 ***	-0.003737 ***	-0.002151 ***	-0.001508 **
	(0.014875)	(0.001004)	(0.000805)	(0.000738)	(0.000858)	(0.000877)	(0.008719)	(0.000908)	(0.000829)	(0.000707)	(0.000692)	(0.000602)
Investments/sales$_{ist-1}$	-0.001661	-0.000181	-0.000122	-0.000107	-0.000195	-0.000207	-0.002000 *	-0.000262 **	-0.000234 *	-0.000199 *	-0.000257 ***	-0.000265 ***
	(0.001739)	(0.000180)	(0.000160)	(0.000141)	(0.000127)	(0.000163)	(0.001153)	(0.000128)	(0.000123)	(0.000101)	(0.000074)	(0.000079)
Log net oper. profit from sales$_{ist-1}$	0.001766	-0.000034	-0.000153	-0.000221	-0.000410	-0.000155	-0.007567	0.000586	0.000271	-0.000051	-0.000538 *	-0.000501 *
	(0.009307)	(0.000452)	(0.000370)	(0.000339)	(0.000374)	(0.000442)	(0.004606)	(0.000375)	(0.000291)	(0.000313)	(0.000289)	(0.000265)
Export/sales$_{ist-1}$	-0.000523	0.000077 ***	0.000079 ***	0.000062 ***	0.000049 ***	0.000032 **	-0.000349	0.000073 ***	0.000063 ***	0.000077 ***	0.000060 ***	0.000044 ***
	(0.000413)	(0.000024)	(0.000019)	(0.000018)	(0.000018)	(0.000015)	(0.000329)	(0.000020)	(0.000022)	(0.000018)	(0.000016)	(0.000014)
Import/sales$_{ist-1}$	-0.000161 **	-0.000043 ***	-0.000044 ***	-0.000037 ***	-0.000029 ***	-0.000025 ***	-0.000100	-0.000053 ***	-0.000047 ***	-0.000046 ***	-0.000036 ***	-0.000028 ***
	(0.000078)	(0.000014)	(0.000011)	(0.000009)	(0.000009)	(0.000006)	(0.000072)	(0.000013)	(0.000013)	(0.000010)	(0.000008)	(0.000007)
Outward FDI stock/GDP$_{it-1}$	-0.000359	0.000294 ***	0.000332 ***	0.000351 ***	0.000378 ***	0.000453 ***	0.000718	0.000241 ***	0.000357 ***	0.000474 ***	0.000539 ***	0.000577 ***
	(0.000823)	(0.000041)	(0.000032)	(0.000039)	(0.000029)	(0.000051)	(0.000580)	(0.000056)	(0.000045)	(0.000044)	(0.000035)	(0.000047)
Inward FDI stock/GDP$_{it-1}$	0.000026	-0.000086 ***	-0.000101 ***	-0.000101 ***	-0.000118 ***	-0.000150 ***	-0.000412 *	-0.000038	-0.000105 ***	-0.000169 ***	-0.000139 ***	-0.000098 ***
	(0.000525)	(0.000032)	(0.000026)	(0.000023)	(0.000021)	(0.000029)	(0.000220)	(0.000040)	(0.000032)	(0.000039)	(0.000026)	(0.000030)
No. of countries i	20	20	20	20	20	20	20	20	20	20	20	20
No. of sectors s	16	16	16	16	16	16	16	16	16	16	16	16
No. of years t	10	10	10	10	10	10	10	10	10	10	10	10

Standard errors in parentheses. A constant is always included but suppressed in the table. The number of observations in the linear model is 1804.
***, ** and * indicate statistical significance at 1%, 5% and 10% level, respectively.

Table 5.A.3: Estimation Results from Control Function Approach

	Results for Pooled Parameter $\check{\alpha}_{ist}$						Results for Pooled First-decile Parameter $\check{\alpha}^1_{ist}$					
		Quantiles Regression Results for Percentile						Quantiles Regression Results for Percentile				
Covariate	Linear	10	25	50	75	90	Linear	10	25	50	75	90
Log patent stock$_{ist-1}$	0.033493 ***	0.021154 ***	0.019954 ***	0.017981 ***	0.015952 ***	0.013393 ***	0.014703 **	0.007535 ***	0.007671 ***	0.007358 ***	0.006334 ***	0.004550 ***
	(0.011816)	(0.002142)	(0.002119)	(0.001872)	(0.002165)	0.002206)	(0.006857)	(0.000958)	(0.000824)	(0.000916)	(0.000840)	(0.000956)
R&D/sales$_{ist-1}$	-0.062753 ***	0.002455 ***	0.002040 ***	0.002118 ***	0.001834 **	0.001840 **	-0.054975 ***	0.000477	0.000581	0.000571 *	0.000729	0.001184 *
	(0.008456)	(0.000935)	(0.000767)	(0.000683)	(0.000780	0.000887)	(0.009554)	(0.000451)	(0.000369)	(0.000344)	(0.000541)	(0.000687)
Share tertiary education$_{it-1}$	-0.004522 ***	0.000111	0.000641	0.000919 **	0.001044) **	0.001359 ***	-0.004337 **	0.000025	0.000193	0.000521 ***	0.000614 ***	0.000778 ***
	(0.001449)	(0.000519)	(0.000469)	(0.000415)	(0.000454	0.000413)	(0.002023)	(0.000157)	(0.000160)	(0.000189)	(0.000186)	(0.000244)
Log no. employees$_{ist-1}$	-0.018358 ***	-0.011212 ***	-0.011431 ***	-0.011666 ***	-0.012070 ***	-0.011604 ***	-0.015366 ***	-0.003452 ***	-0.003727 ***	-0.003637 ***	-0.003845 ***	-0.003589 ***
	(0.003561)	(0.003223)	(0.002639)	(0.002576)	(0.003146)	0.002695)	(0.004342)	(0.001228)	(0.001385)	(0.001199)	(0.000968)	(0.001169)
Investments/sales$_{ist-1}$	-0.000508 *	0.000072	0.000114	0.000094	0.000109)	0.000091	-0.000293	0.000027	0.000051	0.000064	0.000056	-0.000008
	(0.000304)	(0.000539)	(0.000465)	(0.000425)	(0.000418	0.000506)	(0.000254)	(0.000218)	(0.000207)	(0.000216)	(0.000170)	(0.000187)
Log net oper. profit from sales$_{ist-1}$	-0.000246 ***	-0.003914 ***	-0.004043 ***	-0.004318 ***	-0.004301 ***	-0.003983 ***	-0.000342 **	-0.000864	-0.000728	-0.000827	-0.000543	-0.000335
	(0.000055)	(0.001254)	(0.001129)	(0.001126)	(0.001226)	0.001091)	(0.000143)	(0.000568)	(0.000593)	(0.000565)	(0.000490)	(0.000627)
Export/sales$_{ist-1}$	-0.009717	0.000060	0.000060 **	0.000044	0.000037	0.000026	0.002230	0.000018	0.000026 **	0.000020 *	0.000006	-0.000003
	(0.007675)	(0.000044)	(0.000030)	(0.000036)	(0.000032	0.000032)	(0.007126)	(0.000015)	(0.000013)	(0.000012)	(0.000013)	(0.000018)
Import/sales$_{ist-1}$	-0.003502 **	-0.000047 **	-0.000050 ***	-0.000047 ***	-0.000046) ***	-0.000041 **	-0.004119 ***	-0.000014 **	-0.000018 ***	-0.000017 ***	-0.000014 **	-0.000009
	(0.001655)	(0.000020)	(0.000016)	(0.000017)	(0.000017	0.000016)	(0.001598)	(0.000006)	(0.000007)	(0.000006)	(0.000006)	(0.000007)
Outward FDI stock/GDP$_{it-1}$	0.000662	0.001661 ***	0.001583 ***	0.001540 ***	0.001484 ***	0.001456 ***	-0.000285	0.000526 ***	0.000470 ***	0.000459 ***	0.000556 ***	0.000633 ***
	(0.000455)	(0.000148)	(0.000135)	(0.000120)	(0.000133)	0.000122)	(0.000491)	(0.000058)	(0.000044)	(0.000058)	(0.000070)	(0.000079)
Inward FDI stock/GDP$_{it-1}$	0.000169	-0.000718 ***	-0.000710 ***	-0.000685 ***	-0.000660) ***	-0.000673 ***	0.000622	-0.000233 ***	-0.000189 ***	-0.000194 ***	-0.000216 ***	-0.000237 ***
	(0.000473)	(0.000116)	(0.000106)	(0.000112)	(0.000095	0.000107)	(0.000643)	(0.000051)	(0.000034)	(0.000048)	(0.000045)	(0.000074)
No. of countries i	20	20	20	20	20	20	20	20	20	20	20	20
No. of sectors s	16	16	16	16	16	16	16	16	16	16	16	16
No. of years t	8	8	8	8	8	8	8	8	8	8	8	8

	Results for Pooled Fifth-decile Parameter $\check{\alpha}^5_{ist}$						Results for Pooled Ninth-decile Parameter $\check{\alpha}^9_{ist}$					
		Quantiles Regression Results for Percentile						Quantiles Regression Results for Percentile				
Covariate	Linear	10	25	50	75	90	Linear	10	25	50	75	90
Log patent stock$_{ist-1}$	0.048823 ***	0.006566 ***	0.007174 ***	0.007483 ***	0.007732 ***	0.006008 ***	0.013198 *	0.006021 ***	0.006021 ***	0.005141 ***	0.003895 ***	0.002813 ***
	(0.014104)	(0.000907)	(0.000666)	(0.000788)	(0.000756)	(0.001011)	(0.007190)	(0.000928)	(0.000775)	(0.000760)	(0.000798)	(0.000583)
R&D/sales$_{ist-1}$	-0.089929 ***	0.000222	-0.000078	-0.000329	-0.000476	-0.000543	-0.080601 ***	0.000428	0.000537	0.000392	0.000259	0.000325
	(0.007993)	(0.000331)	(0.000287)	(0.000324)	(0.000375)	(0.000451)	(0.006380)	(0.000355)	(0.000369)	(0.000370)	(0.000443)	(0.000472)
Share tertiary education$_{it-1}$	-0.009817 ***	-0.000243 *	-0.000122	0.000087	0.000302 *	0.000483 **	-0.012991 ***	-0.000332 **	-0.000194	-0.000010	-0.000191	-0.000348 *
	(0.002063)	(0.000124)	(0.000145)	(0.000151)	(0.000164)	(0.000198)	(0.001570)	(0.000165)	(0.000156)	(0.000131)	(0.000165)	(0.000181)
Log no. employees$_{ist-1}$	-0.000201	-0.003721 ***	-0.004057 ***	-0.004052 ***	-0.004422 ***	-0.004209 ***	0.005967	-0.005013 ***	-0.004290 ***	-0.003121 ***	-0.001495 *	-0.001062 *
	(0.005605)	(0.001121)	(0.000886)	(0.000902)	(0.001077)	(0.001197)	(0.004325)	(0.001075)	(0.000992)	(0.000841)	(0.000821)	(0.000589)
Investments/sales$_{ist-1}$	-0.000906 ***	-0.000097	-0.000105	-0.000090	-0.000129	-0.000154	-0.000090	-0.000365	-0.000207	-0.000165	-0.000375 **	-0.000492 ***
	(0.000350)	(0.000201)	(0.000181)	(0.000175)	(0.000200)	(0.000180)	(0.000329)	(0.000310)	(0.000207)	(0.000177)	(0.000175)	(0.000151)
Log net oper. profit from sales$_{ist-1}$	-0.000189 ***	0.000522	0.000399	0.000088	-0.000114	-0.000018	-0.000150 ***	0.001302 **	0.000719	0.000258	-0.000403	-0.000281
	(0.000045)	(0.000468)	(0.000393)	(0.000387)	(0.000531)	(0.000536)	(0.000047)	(0.000508)	(0.000474)	(0.000392)	(0.000415)	(0.000344)
Export/sales$_{ist-1}$	-0.032533 **	0.000085 ***	0.000088 ***	0.000076 ***	0.000059 ***	0.000046 **	-0.025121 ***	0.000081 **	0.000074 *	0.000099 ***	0.000083 ***	0.000048 **
	(0.015930)	(0.000024)	(0.000025)	(0.000018)	(0.000018)	(0.000021)	(0.009311)	(0.000034)	(0.000039)	(0.000029)	(0.000016)	(0.000019)
Import/sales$_{ist-1}$	-0.004910 *	-0.000044 ***	-0.000046 ***	-0.000040 ***	-0.000033 ***	-0.000029 ***	-0.003310 ***	-0.000045 **	-0.000041 *	-0.000053 ***	-0.000044 ***	-0.000029 ***
	(0.002528)	(0.000013)	(0.000013)	(0.000008)	(0.000008)	(0.000009)	(0.000842)	(0.000022)	(0.000021)	(0.000015)	(0.000008)	(0.000009)
Outward FDI stock/GDP$_{it-1}$	-0.000460	0.000407 ***	0.000425 ***	0.000440 ***	0.000484 ***	0.000555 ***	-0.000956 ***	0.000383 ***	0.000461 ***	0.000579 ***	0.000627 ***	0.000636 ***
	(0.000419)	(0.000055)	(0.000051)	(0.000050)	(0.000057)	(0.000059)	(0.000181)	(0.000068)	(0.000059)	(0.000056)	(0.000059)	(0.000065)
Inward FDI stock/GDP$_{it-1}$	0.000093	-0.000172 ***	-0.000179 ***	-0.000186 ***	-0.000200 ***	-0.000253 ***	0.001064 ***	-0.000182 ***	-0.000217 ***	-0.000287 ***	-0.000242 ***	-0.000192 ***
	(0.000435)	(0.000045)	(0.000036)	(0.000037)	(0.000032)	(0.000038)	(0.000193)	(0.000066)	(0.000050)	(0.000055)	(0.000041)	(0.000039)
No. of countries i	20	20	20	20	20	20	20	20	20	20	20	20
No. of sectors s	16	16	16	16	16	16	16	16	16	16	16	16
No. of years t	8	8	8	8	8	8	8	8	8	8	8	8

Standard errors in parentheses. Cluster-robust standard errors in the linear model. A constant is always included but suppressed in the table. The number of observations in the linear model is 1370.
Following Terza et al. (2008) and Wooldridge (2015), we use a Control Function Approach to address potential endogeneity.
***, ** and * indicate statistical significance at 1%, 5% and 10% level, respectively.

Table 5.A.4: Estimation Results, Shape Parameter estimated with Generalized Linear Model

	Results for Pooled Parameter $\check{\alpha}_{ist}$					
		Quantiles Regression Results for Percentile				
Covariate	Linear	10	25	50	75	90
Log patent stock$_{ist-1}$	0.020155	0.039087 ***	0.035086 ***	0.030321 ***	0.025557 ***	0.018900 ***
	(0.012269)	(0.003739)	(0.003356)	(0.003232)	(0.002867)	(0.002564)
R&D/sales$_{ist-1}$	-0.002617	0.002132	0.002944 ***	0.003730 ***	0.003764 ***	0.003241 ***
	(0.007174)	(0.001366)	(0.001058)	(0.001095)	(0.001071)	(0.001087)
Share tertiary education$_{it-1}$	0.000607	0.004209 ***	0.003943 ***	0.003863 ***	0.003879 ***	0.003429 ***
	(0.000900)	(0.000844)	(0.000651)	(0.000549)	(0.000583)	(0.000633)
Log no. employees$_{ist-1}$	-0.067550 ***	-0.025142 ***	-0.024939 ***	-0.024638 ***	-0.025574 ***	-0.024004 ***
	(0.012070)	(0.004893)	(0.004754)	(0.004234)	(0.003934)	(0.003903)
Investments/sales$_{ist-1}$	-0.003838 ***	-0.001064	-0.000555	-0.000236	-0.000548	-0.000818
	(0.001054)	(0.001000)	(0.000650)	(0.000643)	(0.000543)	(0.000552)
Log net oper. profit from sales$_{ist-1}$	-0.018642 ***	-0.008241 ***	-0.007163 ***	-0.007120 ***	-0.005300 ***	-0.003282 **
	(0.005626)	(0.002553)	(0.001835)	(0.001859)	(0.001688)	(0.001599)
Export/sales$_{ist-1}$	-0.000016	0.000069	-0.000006	-0.000023	-0.000059	-0.000069 *
	(0.000192)	(0.000106)	(0.000064)	(0.000055)	(0.000059)	(0.000040)
Import/sales$_{ist-1}$	-0.000109	-0.000094	-0.000055 *	-0.000041	-0.000028	-0.000019
	(0.000107)	(0.000059)	(0.000031)	(0.000025)	(0.000027)	(0.000020)
Outward FDI stock/GDP$_{it-1}$	0.000010	0.002165 ***	0.001919 ***	0.001705 ***	0.001422 ***	0.001222 ***
	(0.000409)	(0.000257)	(0.000177)	(0.000182)	(0.000170)	(0.000171)
Inward FDI stock/GDP$_{it-1}$	0.000820 ***	-0.001225 ***	-0.000972 ***	-0.000835 ***	-0.000619 ***	-0.000476 ***
	(0.000296)	(0.000228)	(0.000145)	(0.000160)	(0.000122)	(0.000133)
No. of countries i	20	20	20	20	20	20
No. of sectors s	16	16	16	16	16	16
No. of years t	10	10	10	10	10	10

Standard errors in parentheses. A constant is always included but suppressed in the table. The number of observations in the linear model is 1804.
***, ** and * indicate statistical significance at 1%, 5% and 10% level, respectively.

Table 5.A.5: Estimation Results, Shape Parameter estimated with Generalized Linear Model and Control Function Approach

	Results for Pooled Parameter $\check{\alpha}_{ist}$					
		Quantiles Regression Results for Percentile				
Covariate	Linear	10	25	50	75	90
Log patent stock$_{ist-1}$	0.022116	0.031304 ***	0.027311 ***	0.023383 ***	0.017470 ***	0.011712 ***
	(0.015105)	(0.003692)	(0.003573)	(0.003757)	(0.003418)	(0.004244)
R&D/sales$_{ist-1}$	-0.043528 ***	0.004027 ***	0.004558 ***	0.004328 ***	0.004132 ***	0.003684 ***
	(0.012474)	(0.001533)	(0.001099)	(0.001216)	(0.001062)	(0.001203)
Share tertiary education$_{it-1}$	-0.003292 *	0.002877 ***	0.003221 ***	0.003562 ***	0.003735 ***	0.003413 ***
	(0.001830)	(0.000950)	(0.000702)	(0.000648)	(0.000704)	(0.000761)
Log no. employees$_{ist-1}$	-0.027630 ***	-0.016869 ***	-0.016053 ***	-0.015339 ***	-0.014525 ***	-0.014728 ***
	(0.005796)	(0.004964)	(0.005095)	(0.004283)	(0.004568)	(0.004964)
Investments/sales$_{ist-1}$	-0.000006	0.000535	0.000891	0.000843	0.000736	0.000762
	(0.000280)	(0.000946)	(0.000663)	(0.000573)	(0.000665)	(0.000630)
Log net oper. profit from sales$_{ist-1}$	-0.000358 ***	-0.007746 ***	-0.008066 ***	-0.007985 ***	-0.007752 ***	-0.006678 ***
	(0.000115)	(0.002443)	(0.001961)	(0.002154)	(0.002067)	(0.002125)
Export/sales$_{ist-1}$	-0.000820	-0.000017	-0.000021	-0.000038	-0.000055	-0.000033
	(0.010019)	(0.000079)	(0.000063)	(0.000055)	(0.000051)	(0.000043)
Import/sales$_{ist-1}$	-0.003231	-0.000026	-0.000024	-0.000018	-0.000008	-0.000018
	(0.002111)	(0.000032)	(0.000024)	(0.000023)	(0.000023)	(0.000023)
Outward FDI stock/GDP$_{it-1}$	0.000200	0.002850 ***	0.002471 ***	0.002270 ***	0.002050 ***	0.001932 ***
	(0.000431)	(0.000261)	(0.000225)	(0.000207)	(0.000231)	(0.000243)
Inward FDI stock/GDP$_{it-1}$	0.001201 **	-0.001494 ***	-0.001246 ***	-0.001131 ***	-0.001068 ***	-0.001055 ***
	(0.000497)	(0.000205)	(0.000183)	(0.000171)	(0.000166)	(0.000182)
No. of countries i	20	20	20	20	20	20
No. of sectors s	16	16	16	16	16	16
No. of years t	8	8	8	8	8	8

Standard errors in parentheses. Cluster-robust standard errors in the linear model. A constant is always included but suppressed in the table.
The number of observations in the linear model is 1370.
Following Terza et al. (2008) and Wooldridge (2015), we use a Control Function Approach to address potential endogeneity.
***, ** and * indicate statistical significance at 1%, 5% and 10% level, respectively.

5.B Further Tables

Table 5.B.1: Correlations

	1	2	3	4	5	6	7	8	9	10
1 Log patent stock$_{ist-1}$	1									
2 R&D/sales$_{ist-1}$	0.2999	1								
3 Share tertiary education$_{it-1}$	0.0774	0.2219	1							
4 Log no. employees$_{ist-1}$	0.5717	-0.0106	-0.389	1						
5 Investments/sales$_{ist-1}$	-0.0383	-0.1311	-0.1072	0.0271	1					
6 Log net oper. profit from sales$_{ist-1}$	0.2549	0.0063	-0.2279	0.4525	0.2861	1				
7 Export/sales$_{ist-1}$	0.0416	0.1322	0.178	-0.2054	-0.1571	-0.1856	1			
8 Import/sales$_{ist-1}$	0.0634	0.0168	0.1228	-0.269	-0.0333	-0.1146	0.7924	1		
9 Outward FDI stock/GDP$_{it-1}$	0.185	0.1457	0.649	-0.2309	-0.1791	-0.2872	0.2354	0.0989	1	
10 Inward FDI stock/GDP$_{it-1}$	-0.1939	-0.0369	0.3457	-0.2878	-0.107	-0.1513	0.2515	0.1207	0.6012	1

Bibliography

Acharya, Ram C (2015). "Revisiting measure of R&D spillovers: empirical evidence on OECD countries and industries." In: *Economics of Innovation and New Technology* 24 (4), pp. 360–400.

Aghion, Philippe and Xavier Jaravel (2015). "Knowledge spillovers, innovation and growth." In: *The Economic Journal* 125 (583), pp. 533–573.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt (2005). "Competition and Innovation: An Inverted-U Relationship." In: *The Quarterly Journal of Economics* 120 (2), pp. 701–728.

Aghion, Philippe, Richard Blundell, Rachel Griffith, Peter Howitt, and Susanne Prantl (2009). "The effects of entry on incumbent innovation and productivity." In: *The Review of Economics and Statistics* 91 (1), pp. 20–32.

Aghion, Philippe, Antoine Dechezleprêtre, David Hemous, Ralf Martin, and John Van Reenen (2012). "Carbon taxes, path dependency and directed technical change: evidence from the auto industry." In: *NBER Working Paper No. 18596*.

Arellano, Manuel and Stephen Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." In: *The Review of Economic Studies* 58 (2), pp. 277–297.

Arenallo, Manuel and Olympia Bover (1995). "Another look at the instrumental variable estimation of error-components models." In: *Journal of Econometrics* 68 (1), pp. 29–51.

Arora, Ashish, Alfonso Gambardella, Laura Magazzini, and Fabio Pammolli (2009). "A breath of fresh air? Firm type, scale, scope, and selection effects in drug development." In: *Management Science* 55 (10), pp. 1638–1653.

Arvanitis, Spyros (1997). "The Impact of Firm Size on Innovative Activity. An Empirical Analysis Based on Swiss Firm Data." In: *Small Business Economics* 9 (6), pp. 473–490.

Arvanitis, Spyros and Heinz Hollenstein (1996). “Industrial Innovation in Switzerland: A Model-based Analysis with Survey Data.” In: *Kleinknecht, A. (ed.): Determinants of Innovation. The Message from New Indicators, Macmillan*, pp. 13–62.

Aschhoff, Birgit and Tobias Schmidt (2008). “Empirical evidence on the success of R&D cooperation – Happy together?” In: *Review of Industrial Organization* 33 (1), pp. 41–62.

Audretsch, David B and Maryann P Feldman (1996). “R&D spillovers and the geography of innovation and production.” In: *The American Economic Review* 86 (3), pp. 630–640.

Bacchiocchi, Emanuele and Fabio Montobbio (2010). “International Knowledge Diffusion and Home-bias Effect: Do USPTO and EPO Patent Citations Tell the Same Story?” In: *The Scandinavian Journal of Economics* 112 (3), pp. 441–470.

Badinger, Harald and Peter Egger (2015). “Productivity Spillovers Across Countries and Industries: New Evidence From OECD Countries.” In: *Oxford Bulletin of Economics and Statistics* 78 (4), pp. 501–521.

Baltagi, Badi (2008). *Econometric analysis of panel data.* John Wiley & Sons.

Belderbos, René, Martin Carree, and Boris Lokshin (2004). “Cooperative R&D and firm performance.” In: *Research policy* 33 (10), pp. 1477–1492.

Belenzon, Sharon (2012). “Cumulative Innovation and Market Value: Evidence from Patent Citations.” In: *The Economic Journal* 122 (559), pp. 265–285.

Blindenbach-Driessen, Floortje and Jan Ende (2014). “The locus of innovation: The effect of a separate innovation unit on exploration, exploitation, and ambidexterity in manufacturing and service firms.” In: *Journal of Product Innovation Management* 31 (5), pp. 1089–1105.

Bloch, Carter (2013). “R&D spillovers and productivity: an analysis of geographical and technological dimensions.” In: *Economics of Innovation and New Technology* 22 (5), pp. 447–460.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013). “Identifying technology spillovers and product market rivalry.” In: *Econometrica* 81 (4), pp. 1347–1393.

Blundell, Richard and Stephen Bond (1998). “Initial conditions and moment restrictions in dynamic panel data models.” In: *Journal of Econometrics* 87 (1), pp. 115–143.

Boesenberg, Simon and Peter H. Egger (2016). “R&D tax incentives and the emergence and trade of ideas.” In: *Economic Policy, 63rd Panel Meeting, Amsterdam.*

Breschi, Stefano, Francesco Lissoni, and Franco Malerba (2003). “Knowledge-relatedness in firm technological diversification.” In: *Research Policy* 32 (1), pp. 69–87.

Bstieler, Ludwig (2005). “The moderating effect of environmental uncertainty on new product development and time efficiency.” In: *Journal of Product Innovation Management* 22 (3), pp. 267–284.

Calantone, Roger, Rosanna Garcia, and Cornelia Dröge (2003). “The effects of environmental turbulence on new product development strategy planning.” In: *Journal of Product Innovation Management* 20 (2), pp. 90–103.

Calantone, Roger and Gaia Rubera (2012). “When should RD&E and marketing collaborate? The moderating role of exploration–exploitation and environmental uncertainty.” In: *Journal of Product Innovation Management* 29 (1), pp. 144–157.

Cameron, A Colin and Pravin K Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge university press.

Cameron, A Colin and Pravin K Trivedi (2013). *Regression Analysis of Count Data*. Vol. 53. Cambridge university press.

Canay, Ivan A (2011). “A simple approach to quantile regression for panel data.” In: *The Econometrics Journal* 14 (3), pp. 368–386.

Cappelli, Riccardo, Dirk Czarnitzki, and Kornelius Kraft (2014). “Sources of spillovers for imitation and innovation.” In: *Research Policy* 43 (1), pp. 115–120.

Cappelli, Riccardo and Fabio Montobbio (2014). “European Integration and Knowledge Flows across European Regions.” In: *Regional Studies* 50 (4), pp. 1–19.

Cardamone, Paola (2010). “The role of R&D spillovers in product and process innovation.” In: *Applied Economics Letters* 17 (5), pp. 483–493.

Cassiman, Bruno and Reinhilde Veugelers (2002). “R&D cooperation and spillovers: some empirical evidence from Belgium.” In: *The American Economic Review* 92 (4), pp. 1169–1184.

Cassiman, Bruno and Reinhilde Veugelers (2006). “In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition.” In: *Management science* 52 (1), pp. 68–82.

Chandy, Rajesh K and Gerard J Tellis (2000). “The incumbents curse? Incumbency, size, and radical product innovation.” In: *Journal of marketing* 64 (3), pp. 1–17.

Chen, Eric L and Riitta Katila (2008). “Rival interpretations of balancing exploration and exploitation: simultaneous or sequential.” In: *Handbook of Technology and Innovation Management*, pp. 197–214.

Chernozhukov, Victor and Christian Hansen (2008). "Instrumental variable quantile regression: A robust inference approach." In: *Journal of Econometrics* 142 (1), pp. 379–398.

Chesbrough, Henry William (2006). *Open innovation: The new imperative for creating and profiting from technology.* Harvard Business Press.

Chiu, Yi-Chia, Hsien-Che Lai, Tai-Yu Lee, and Yi-Ching Liaw (2008). "Technological diversification, complementary assets, and performance." In: *Technological Forecasting and Social Change* 75 (6), pp. 875–892.

Cockburn, Iain and Zvi Griliches (1988). "Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents." In: *The American Economic Review* 78 (2), pp. 419–423.

Cockburn, Iain M and Rebecca M Henderson (2001). "Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research." In: *Journal of Health Economics* 20 (6), pp. 1033–1057.

Coe, David T and Elhanan Helpman (1995). "International R&D spillovers." In: *European Economic Review* 39 (5), pp. 859–887.

Cohen, Wesley M (2010). "Fifty years of empirical studies of innovative activity and performance." In: *Hall, B.H. and N. Rosenberg (ed.): Handbook of the Economics of Innovation* 1, pp. 129–213.

Cohen, Wesley M and Daniel A Levinthal (1989). "Innovation and learning: the two faces of R & D." In: *The Economic Journal* 99 (397), pp. 569–596.

Cohen, Wesley M and Daniel A Levinthal (1990). "Absorptive capacity: A new perspective on learning and innovation." In: *Administrative Science Quarterly* 35 (1), pp. 128–152.

Cohen, Wesley M, Akira Goto, Akiya Nagata, Richard R Nelson, and John P Walsh (2002). "R&D spillovers, patents and the incentives to innovate in Japan and the United States." In: *Research Policy* 31 (8), pp. 1349–1367.

Crépon, Bruno, Emmanuel Duguet, and Jacques Mairessec (1998). "Research, Innovation And Productivity: An Econometric Analysis At The Firm Level." In: *Economics of Innovation and New Technology* 7 (2), pp. 115–158.

Criscuolo, Paola and Bart Verspagen (2008). "Does it matter where patent citations come from? Inventor vs. examiner citations in European patents." In: *Research Policy* 37 (10), pp. 1892–1908.

Czarnitzki, Dirk and Susanne Thorwarth (2012). "Productivity effects of basic research in low-tech and high-tech industries." In: *Research Policy* 41 (9), pp. 1555–1564.

Dahlin, Kristina B and Dean M Behrens (2005). "When is an invention really radical?: Defining and measuring technological radicalness." In: *Research Policy* 34 (5), pp. 717–737.

Danneels, Erwin (2002). "The Dynamics of Product Innovation and Firm Competences." In: *Strategic Management Journal* 23 (12), pp. 1095–1121.

Danzon, Patricia M, Sean Nicholson, and Nuno Sousa Pereira (2005). "Productivity in pharmaceutical–biotechnology R&D: the role of experience and alliances." In: *Journal of Health Economics* 24 (2), pp. 317–339.

Dawson, Jeremy (2012). "Measurement of work group diversity." PhD thesis. Aston University.

De Bondt, Raymond (1997). "Spillovers and innovative activities." In: *International Journal of Industrial Organization* 15 (1), pp. 1–28.

De La Potterie, Bruno Van Pottelsberghe (1997). "Issues in assessing the effect of interindustry R&D spillovers." In: *Economic Systems Research* 9 (4), pp. 331–356.

De Rassenfosse, Gaétan, Hélène Dernis, and Geert Boedt (2014). "An introduction to the Patstat database with example queries." In: *Australian Economic Review* 47 (3), pp. 395–408.

De Rassenfosse, Gaétan, Hélène Dernis, Dominique Guellec, Lucio Picci, and Bruno van Pottelsberghe de la Potterie (2013). "The worldwide count of priority patents: A new indicator of inventive activity." In: *Research Policy* 42 (3), pp. 720–737.

Dhont-Peltrault, Estelle and Etienne Pfister (2011). "R&D cooperation versus R&D subcontracting: empirical evidence from French survey data." In: *Economics of Innovation and New Technology* 20 (4), pp. 309–341.

Dietzenbacher, Erik, Bart Los, Robert Stehrer, Marcel Timmer, and Gaaitzen De Vries (2013). "The construction of world input–output tables in the WIOD project." In: *Economic Systems Research* 25 (1), pp. 71–98.

Du Plessis, Mariette, Bart Van Looy, Xiaoyan Song, and Tom Magerman (2009). "Data production methods for harmonized patent statistics: Patent sector allocation." In: *Eurostat Working Paper*.

Duguet, Emmanuel and Megan MacGarvie (2005). "How well do patent citations measure flows of technology? Evidence from French innovation surveys." In: *Economics of Innovation and New Technology* 14 (5), pp. 375–393.

Egger, Peter and Michaela Kesina (2013). “Financial constraints and exports: evidence from Chinese firms.” In: *CESifo Economic Studies* 59 (4), pp. 676–706.

Egger, Peter and Michaela Kesina (2014). “Financial constraints and the extensive and intensive margin of firm exports: Panel data evidence from China.” In: *Review of Development Economics* 18 (4), pp. 625–639.

Egger, Peter and Michael Pfaffermayr (2003). “The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects.” In: *Empirical Economics* 28 (3), pp. 571–580.

Eng, Teck-Yong and Sena Ozdemir (2014). “International R&D partnerships and intrafirm R&D–marketing–production integration of manufacturing firms in emerging economies.” In: *Industrial Marketing Management* 43 (1), pp. 32–44.

EPO, European Patent Office (2013). *PATSTAT Database, October 2013.*

Feldman, Maryann P and Dieter F Kogler (2010). “Stylized facts in the geography of innovation.” In: *Hall, B.H. and N. Rosenberg (ed.): Handbook of the Economics of Innovation* 1, pp. 381–410.

Fischer, Manfred M, Thomas Scherngell, and Eva Jansenberger (2009). “Geographic localisation of knowledge spillovers: evidence from high-tech patent citations in Europe.” In: *The Annals of Regional Science* 43 (4), pp. 839–858.

Fleming, Lee (2001). “Recombinant uncertainty in technological search.” In: *Management Science* 47 (1), pp. 117–132.

Frantzen, Dirk (2002). “Intersectoral and international R&D knowledge spillovers and total factor productivity.” In: *Scottish Journal of Political Economy* 49 (3), pp. 280–303.

Gambardella, Alfonso and Salvatore Torrisi (1998). “Does technological convergence imply convergence in markets? Evidence from the electronics industry.” In: *Research Policy* 27 (5), pp. 445–463.

Garcia, Rosanna and Roger Calantone (2002). “A critical look at technological innovation typology and innovativeness terminology: a literature review.” In: *Journal of Product Innovation Management* 19 (2), pp. 110–132.

Garcia-Vega, Maria (2006). “Does technological diversification promote innovation?: An empirical analysis for European firms.” In: *Research Policy* 35 (2), pp. 230–246.

Garriga, Helena, Georg von Krogh, and Sebastian Spaeth (2013). “How constraints and knowledge impact open innovation.” In: *Strategic Management Journal* 34 (9), pp. 1134–1144.

Gemba, Kiminori and Fumio Kodama (2001). "Diversification dynamics of the Japanese industry." In: *Research Policy* 30 (8), pp. 1165–1184.

Geroski, Paul A (1995). "Do spillovers undermine the incentive to innovate?" In: *Dowrick, P.A. (ed.): Economic Approaches to Innovation.*

Gertler, Meric S (2003). "Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there)." In: *Journal of Economic Geography* 3 (1), pp. 75–99.

Granstrand, Ove (1998). "Towards a theory of the technology-based firm." In: *Research Policy* 27 (5), pp. 465–489.

Granstrand, Ove, Pari Patel, and Keith Pavitt (1997). "Multi-technology corporations: why they have "distributed" rather than "distinctive core" competencies." In: *California Management Review* 39 (4), pp. 8–25.

Griffin, Abbie and John R Hauser (1996). "Integrating R&D and marketing: a review and analysis of the literature." In: *Journal of Product Innovation Management* 13 (3), pp. 191–215.

Griffith, Rachel, Stephen Redding, and John Van Reenen (2004). "Mapping the two faces of R&D: productivity growth in a panel of OECD industries." In: *Review of Economics and Statistics* 86 (4), pp. 883–895.

Griliches, Z (1995). "R&D and productivity: econometric results and productivity issues." In: *Stoneman, P. (ed.): Handbook of the Economics of Innovation and Technological Change, Oxford, Basil Blackwell.*

Griliches, Zvi (1979). "Issues in assessing the contribution of research and development to productivity growth." In: *The Bell Journal of Economics* 10 (1), pp. 92–116.

Griliches, Zvi (1992). "The Search for R&D Spillovers." In: *Scandinavian Journal of Economics* 94, S29–47.

Gupta, Anil K, Ken G Smith, and Christina E Shalley (2006). "The interplay between exploration and exploitation." In: *Academy of Management Journal* 49 (4), pp. 693–706.

Gupta, Ashok K, SP Raj, and David Wilemon (1986). "A model for studying R&D. Marketing interface in the product innovation process." In: *The Journal of Marketing* 50 (2), pp. 7–17.

Gust-Bardon, Natalia Irena et al. (2012). "The role of geographical proximity in innovation: Do regional and local levels really matter?" In: *Working Papers firms and region R4/2012.*

Hacklin, Fredrik, Boris Battistini, and Georg Von Krogh (2013). "Strategic choices in converging industries." In: *MIT Sloan Management Review* 55 (1), pp. 65–73.

Hagedoorn, John and Geert Duysters (2002). "External sources of innovative capabilities: the preferences for strategic alliances or mergers and acquisitions." In: *Journal of Management Studies* 39 (2), pp. 167–188.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg (2005). "Market value and patent citations." In: *RAND Journal of Economics* 36 (1), pp. 16–38.

Hall, Bronwyn H, Jacques Mairesse, and Pierre Mohnen (2010). "Measuring the Returns to R&D." In: *Hall, B.H. and N. Rosenberg (ed.): Handbook of the Economics of Innovation* 2, pp. 1033–1082.

Harhoff, Dietmar, Frederic M Scherer, and Katrin Vopel (2003). "Citations, family size, opposition and the value of patent rights." In: *Research Policy* 32 (8), pp. 1343–1363.

Hatzichronoglou, Thomas (1997). "Revision of the High-Technology Sector and Product Classification." In: *OECD Science, Technology and Industry Working Papers, 1997/02.*

He, Zi-Lin and Poh-Kam Wong (2004). "Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis." In: *Organization Science* 15 (4), pp. 481–494.

Henderson, Rebecca and Iain Cockburn (1996). "Scale, scope, and spillovers: the determinants of research productivity in drug discovery." In: *The Rand Journal of Economics* 27 (1), pp. 32–59.

Henderson, Rebecca, Adam Jaffe, and Manuel Trajtenberg (2005). "Patent citations and the geography of knowledge spillovers: A reassessment: Comment." In: *The American Economic Review* 95 (1), pp. 461–464.

Hill, Susan A and Julian Birkinshaw (2006). "Ambidexterity in corporate venturing: Simultaneously using existing and building new capabilities." In: *Academy of Management Proceedings.* Vol. 2006. 1, pp. C1–C6.

Hu, Albert GZ, Gary H Jefferson, and Qian Jinchang (2005). "R&D and technology transfer: firm-level evidence from Chinese industry." In: *Review of Economics and Statistics* 87 (4), pp. 780–786.

Jacquemin, Alexis P and Charles H Berry (1979). "Entropy measure of diversification and corporate growth." In: *The Journal of Industrial Economics* 27 (4), pp. 359–369.

Jaffe, Adam B (1986). "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value." In: *The American Economic Review* 76 (5), pp. 984–1001.

Jaffe, Adam B (1989). “Characterizing the technological position of firms, with application to quantifying technological opportunity and research spillovers.” In: *Research Policy* 18 (2), pp. 87–97.

Jaffe, Adam B and Manuel Trajtenberg (1999). “International knowledge flows: evidence from patent citations.” In: *Economics of Innovation and New Technology* 8 (1-2), pp. 105–136.

Jaffe, Adam B and Manuel Trajtenberg (2002). *Patents, citations, and innovations: A window on the knowledge economy.* MIT press.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson (1993). “Geographic localization of knowledge spillovers as evidenced by patent citations.” In: *The Quarterly Journal of Economics* 108 (3), pp. 577–598.

Jansen, Justin JP, Frans AJ Van Den Bosch, and Henk W Volberda (2006). “Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators.” In: *Management Science* 52 (11), pp. 1661–1674.

Jirjahn, Uwe and Kornelius Kraft (2011). “Do spillovers stimulate incremental or drastic product innovations? Evidence from German establishment data.” In: *Oxford Bulletin of Economics and Statistics* 73 (4), pp. 509–538.

Katila, Riitta and Gautam Ahuja (2002). “Something old, something new: A longitudinal study of search behavior and new product introduction.” In: *Academy of Management Journal* 45 (6), pp. 1183–1194.

Keller, Wolfgang (2002a). “Geographic localization of international technology diffusion.” In: *The American Economic Review* 92 (1), pp. 120–142.

Keller, Wolfgang (2002b). “Trade and the Transmission of Technology.” In: *Journal of Economic Growth* 7 (1), pp. 5–24.

Keller, Wolfgang (2004). “International Technology Diffusion.” In: *Journal of Economic Literature* 42 (3), pp. 752–782.

Kim, Hyukjoon, Hyojeong Lim, and Yongtae Park (2009). “How should firms carry out technological diversification to improve their performance? An analysis of patenting of Korean firms.” In: *Economics of Innovation and New Technology* 18 (8), pp. 757–770.

Kim, Jungho, Chang-Yang Lee, and Yunok Cho (2016). “Technological diversification, core-technology competence, and firm growth.” In: *Research Policy* 45 (1), pp. 113–124.

Kleinknecht, Alfred, Kees Van Montfort, and Erik Brouwer (2002). "The non-trivial choice between innovation indicators." In: *Economics of Innovation and new technology* 11 (2), pp. 109–121.

Köhler, Christian, Wolfgang Sofka, and Christoph Grimpe (2012). "Selective search, sectoral patterns, and the impact on product innovation performance." In: *Research Policy* 41 (8), pp. 1344–1356.

Lanjouw, Jean O and Mark Schankerman (2004). "Patent quality and research productivity: Measuring innovation with multiple indicators." In: *The Economic Journal* 114 (495), pp. 441–465.

Laursen, Keld and Ammon Salter (2006). "Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms." In: *Strategic management journal* 27 (2), pp. 131–150.

Lavie, Dovev and Lori Rosenkopf (2006). "Balancing exploration and exploitation in alliance formation." In: *Academy of Management Journal* 49 (4), pp. 797–818.

Lavie, Dovev, Uriel Stettner, and Michael L Tushman (2010). "Exploration and exploitation within and across organizations." In: *The Academy of Management Annals* 4 (1), pp. 109–155.

Leten, Bart, Rene Belderbos, and Bart Van Looy (2007). "Technological diversification, coherence, and performance of firms." In: *Journal of Product Innovation Management* 24 (6), pp. 567–579.

Levinthal, Daniel A and James G March (1993). "The myopia of learning." In: *Strategic Management Journal* 14 (S2), pp. 95–112.

Li, Yao Amber (2014). "Borders and distance in knowledge spillovers: Dying over time or dying with age?Evidence from patent citations." In: *European Economic Review* 71 (0), pp. 152 –172.

Li, Ying, Wim Vanhaverbeke, and Wilfred Schoenmakers (2008). "Exploration and exploitation in innovation: Reframing the interpretation." In: *Creativity and Innovation Management* 17 (2), pp. 107–126.

Lybbert, Travis J and Nikolas J Zolas (2014). "Getting patents and economic data to speak to each other: An algorithmic links with probabilities approach for joint analyses of patenting and economic activity." In: *Research Policy* 43 (3), pp. 530–542.

Lychagin, Sergey, Joris Pinkse, Margaret E Slade, and John Van Reenen (2016). "Spillovers in Space: Does Geography Matter?" In: *The Journal of Industrial Economics* 64 (2), pp. 295–335.

MacGarvie, Megan (2006). "Do firms learn from international trade?" In: *Review of Economics and Statistics* 88 (1), pp. 46–60.

Magerman, Tom, Bart Van Looy, and Xiaoyan Song (2009). "Data production methods for harmonized patent statistics: Patentee name harmonization." In: *Eurostat Working Paper*.

Malerba, Franco, M. L. Mancusi, and Fabio Montobbio (2013). "Innovation, international R&D spillovers and the sectoral heterogeneity of knowledge flows." In: *Review of World Economics* 149 (4), pp. 697–722.

Mancusi, Maria Luisa (2008). "International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations." In: *Journal of International Economics* 76 (2), pp. 155 –165.

March, James G (1991). "Exploration and exploitation in organizational learning." In: *Organization science* 2 (1), pp. 71–87.

Martínez, Catalina (2011). "Patent families: when do different definitions really matter?" In: *Scientometrics* 86 (1), pp. 39–63.

Maurseth, Per Botolf and Bart Verspagen (2002). "Knowledge spillovers in Europe: a patent citations analysis." In: *The Scandinavian Journal of Economics* 104 (4), pp. 531–545.

Mayer, Thierry and Soledad Zignago (2011). *Notes on CEPIIs distances measures: The GeoDist database*. Working Papers 2011-25. CEPII.

McKinsey&Company (2016). *Automotive revolution - perspective towards 2030*.

Miguélez, Ernest and Rosina Moreno (2015). "Knowledge flows and the absorptive capacity of regions." In: *Research Policy* 44 (4), pp. 833–848.

Miller, Douglas J (2006). "Technological diversity, related diversification, and firm performance." In: *Strategic Management Journal* 27 (7), pp. 601–619.

Miotti, Luis and Frédérique Sachwald (2003). "Co-operative R&D: why and with whom?: An integrated framework of analysis." In: *Research policy* 32 (8), pp. 1481–1499.

Morescalchi, Andrea, Fabio Pammolli, Orion Penner, Alexander M Petersen, and Massimo Riccaboni (2014). "The evolution of networks of innovators within and across borders: Evidence from patent data." In: *Research Policy* 44 (3), pp. 651–668.

Mukherji, Nivedita and Jonathan Silberman (2013). "Absorptive capacity, knowledge flows, and innovation in US metropolitan areas." In: *Journal of Regional Science* 53 (3), pp. 392–417.

Murata, Yasusada, Ryo Nakajima, Ryosuke Okamoto, and Ryuichi Tamura (2014). "Localized knowledge spillovers and patent citations: A distance-based approach." In: *Review of Economics and Statistics* 96 (5), pp. 967–985.

Nelson, Richard R and Sidney G Winter (1982). *An evolutionary theory of economic change.* Harvard University Press.

Nerkar, Atul and Peter W Roberts (2004). "Technological and product-market experience and the success of new product introductions in the pharmaceutical industry." In: *Strategic Management Journal* 25 (8-9), pp. 779–799.

Nesta, Lionel and Pier Paolo Saviotti (2005). "Coherence of the knowledge base and the firm's innovative performance: evidence from the US pharmaceutical industry." In: *The Journal of Industrial Economics* 53 (1), pp. 123–142.

OECD (2005). *The OECD STAN database for Industrial Analysis.* Paris.

OECD (2009). *OECD patent statistics manual.* Paris.

OECD (2012). *The OECD Analytical BERD (ANBERD) database.* Paris.

O'Neale, Dion RJ and Shaun C Hendy (2012). "Power law distributions of patents as indicators of innovation." In: *PloS One* 7 (12), e49501.

Paci, Raffaele, Emanuela Marrocu, and Stefano Usai (2014). "The complementary effects of proximity dimensions on knowledge spillovers." In: *Spatial Economic Analysis* 9 (1), pp. 9–30.

Pakes, Ariel (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." In: *Econometrica* 54 (4), pp. 755–784.

Pakes, Ariel and Zvi Griliches (1984). "Estimating distributed lags in short panels with an application to the specification of depreciation patterns and capital stock constructs." In: *The Review of Economic Studies* 51 (2), pp. 243–262.

Papke, Leslie E and Jeffrey M Wooldridge (1996). "Econometric methods for fractional response variables with an application to 401 (k) plan participation rates." In: *Journal of Applied Econometrics* 11 (6), pp. 619–632.

Park, Jungsoo (2004). "International and intersectoral R&D spillovers in the OECD and East Asian economies." In: *Economic Inquiry* 42 (4), pp. 739–757.

Pavitt, Keith (1998). "Technologies, products and organization in the innovating firm: what Adam Smith tells us and Joseph Schumpeter doesn't." In: *Industrial and Corporate Change* 7 (3), pp. 433–452.

Peeters, Bert, Xiaoyan Song, Julie Callaert, Grouwels Joris, and Bart Van Looy (2009). “Harmonizing harmonized patentee nname: An exploratory assessment of top patentees.” In: *Eurostat Working Paper*.

Peri, Giovanni (2005). “Determinants of knowledge flows and their effect on innovation.” In: *Review of Economics and Statistics* 87 (2), pp. 308–322.

Plotnikova, Tatiana et al. (2010). “Success in pharmaceutical research: The changing role of scale and scope economies, spillovers and competition.” In: *Jena Economic Research Papers* 45.

Quintana-García, Cristina and Carlos A Benavides-Velasco (2008). “Innovative competence, exploration and exploitation: The influence of technological diversification.” In: *Research Policy* 37 (3), pp. 492–507.

Raisch, Sebastian, Julian Birkinshaw, Gilbert Probst, and Michael L Tushman (2009). “Organizational ambidexterity: Balancing exploitation and exploration for sustained performance.” In: *Organization Science* 20 (4), pp. 685–695.

Ramani, Shyama V, Mhamed-Ali El-Aroui, and Myriam Carrère (2008). “On estimating a knowledge production function at the firm and sector level using patent statistics.” In: *Research Policy* 37 (9), pp. 1568–1578.

Roach, Michael and Wesley M Cohen (2013). “Lens or prism? Patent citations as a measure of knowledge flows from public research.” In: *Management Science* 59 (2), pp. 504–525.

Romer, Paul M (1990). “Endogenous Technological Change.” In: *Journal of Political Economy* 98 (5 pt 2), pp. 71–102.

Rothaermel, Frank T (2001). “Incumbent’s advantage through exploiting complementary assets via interfirm cooperation.” In: *Strategic Management Journal* 22 (6-7), pp. 687–699.

Rothaermel, Frank T and David L Deeds (2004). “Exploration and exploitation alliances in biotechnology: A system of new product development.” In: *Strategic Management Journal* 25 (3), pp. 201–221.

Ruekert, Robert W and Orville C Walker Jr (1987). “Marketing’s interaction with other functional units: A conceptual framework and empirical evidence.” In: *The Journal of Marketing* 51 (1), pp. 1–19.

Saxonhouse, Gary R (1977). “Regressions from samples having different characteristics.” In: *The Review of Economics and Statistics* 59 (2), pp. 234–237.

Scherer, Frederic (1984). "Using linked patent and R&D data to measure interindustry technology flows." In: *Griliches, Z. (ed.): R&D, patents, and productivity.* University of Chicago Press, pp. 417–464.

Scherer, Frederic M (1965). "Firm size, market structure, opportunity, and the output of patented inventions." In: *The American Economic Review* 55 (5), pp. 1097–1125.

Scherer, Frederic M (1967). "Market structure and the employment of scientists and engineers." In: *The American Economic Review* 57 (3), pp. 524–531.

Scherer, Frederic M (2000). "The size distribution of profits from innovation." In: *The Economics and Econometrics of Innovation.* Springer, pp. 473–494.

Scherer, Frederic M and Dietmar Harhoff (2000). "Technology policy for a world of skew-distributed outcomes." In: *Research Policy* 29 (4), pp. 559–566.

Schoenmakers, Wilfred and Geert Duysters (2010). "The technological origins of radical inventions." In: *Research Policy* 39 (8), pp. 1051–1059.

Seliger, Florian and Spyros Arvanitis (2014). "Imitation vs Innovation-What Makes the Difference?" In: *KOF Working Papers No. 367.*

Silverberg, Gerald and Bart Verspagen (2007). "The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance." In: *Journal of Econometrics* 139 (2), pp. 318–339.

Simsek, Zeki (2009). "Organizational ambidexterity: Towards a multilevel understanding." In: *Journal of Management Studies* 46 (4), pp. 597–624.

Singh, Kulwant (1997). "The impact of technological complexity and interfirm cooperation on business survival." In: *Academy of Management Journal* 40 (2), pp. 339–367.

Song, Michael and R Jeffrey Thieme (2006). "A cross-national investigation of the R&D–marketing interface in the product innovation process." In: *Industrial Marketing Management* 35 (3), pp. 308–322.

Souder, William E (1981). "Disharmony between R&D and marketing." In: *Industrial Marketing Management* 10 (1), pp. 67–73.

Souder, William E (1988). "Managing relations between R&D and marketing in new product development projects." In: *Journal of Product Innovation Management* 5 (1), pp. 6–19.

Spence, Michael (1984). "Cost reduction, competition, and industry performance." In: *Econometrica: Journal of the Econometric Society* 52 (1), pp. 101–121.

Squicciarini, Mariagrazia, Hélène Dernis, and Chiara Criscuolo (2013). "Measuring patent quality: Indicators of technological and economic value." In: *OECD Science, Technology and Industry Working Papers, 2013/03.*

Stirling, Andrew (1998). "On the economics and analysis of diversity." In: *Science Policy Research Unit (SPRU), Electronic Working Papers Series, Paper* 28, pp. 1–156.

Terza, Joseph V, Anirban Basu, and Paul J Rathouz (2008). "Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling." In: *Journal of health economics* 27 (3), pp. 531–543.

Tether, Bruce S (2002). "Who co-operates for innovation, and why: an empirical analysis." In: *Research policy* 31 (6), pp. 947–967.

The World Bank (2014). *World Development Indicators.*

Thompson, Peter and Melanie Fox-Kean (2005). "Patent citations and the geography of knowledge spillovers: A reassessment." In: *American Economic Review* 95 (1), pp. 450–460.

Toh, Puay Khoon and Taekyu Kim (2013). "Why put all your eggs in one basket? A competition-based view of how technological uncertainty affects a firms technological specialization." In: *Organization Science* 24 (4), pp. 1214–1236.

Torrisi, Salvatore and Ove Grandstrand (2004). "Technological and business diversification - A survey of theories and empirical evidence." In: *Cantwell, John and Gambardella, Alfonso and Granstrand, Ove (eds.): The Economics and Management of Technological Diversification.*

Trajtenberg, Manuel (1990). "A penny for your quotes: patent citations and the value of innovations." In: *The Rand Journal of Economics* 21 (1), pp. 172–187.

Tsai, Kuen-Hung and Jiann-Chyuan Wang (2004). "R&D Productivity and the Spillover Effects of High-tech Industry on the Traditional Manufacturing Sector: The Case of Taiwan." In: *The World Economy* 27 (10), pp. 1555–1570.

Uotila, Juha, Markku Maula, Thomas Keil, and Shaker A Zahra (2009). "Exploration, exploitation, and financial performance: analysis of S&P 500 corporations." In: *Strategic Management Journal* 30 (2), pp. 221–231.

Verspagen, Bart (1997). "Estimating international technology spillovers using technology flow matrices." In: *Weltwirtschaftliches Archiv* 133 (2), pp. 226–248.

Verspagen, Bart and Ivo De Loo (1999). "Technology spillovers between sectors." In: *Technological Forecasting and Social Change* 60 (3), pp. 215–235.

Von Hippel, Eric (2005). "Democratizing innovation: The evolving phenomenon of user innovation." In: *Journal für Betriebswirtschaft* 55 (1), pp. 63–78.

Weitzman, Martin L (1998). "Recombinant growth." In: *Quarterly journal of Economics* 113 (2), pp. 331–360.

Wersching, Klaus (2010). "Schumpeterian competition, technological regimes and learning through knowledge spillover." In: *Journal of Economic Behavior & Organization* 75 (3), pp. 482–493.

West, Joel and Marcel Bogers (2014). "Leveraging external sources of innovation: a review of research on open innovation." In: *Journal of Product Innovation Management* 31 (4), pp. 814–831.

West, Joel, Ammon Salter, Wim Vanhaverbeke, and Henry Chesbrough (2014). "Open innovation: The next decade." In: *Research Policy* 43 (5), pp. 805–811.

WIPO, World Intellectual Property Organization (2014). *International patent classification - guide*. Geneva.

Wolff, Edward N and M Ishaq Nadiri (1993). "Spillover effects, linkage structure, and research and development." In: *Structural Change and Economic Dynamics* 4 (2), pp. 315–331.

Wooldridge, Jeffrey M (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, Jeffrey M (2015). "Control function methods in applied econometrics." In: *Journal of Human Resources* 50 (2), pp. 420–445.

Yang, Hongyan, Corey Phelps, and H Kevin Steensma (2010). "Learning from what others have learned from you: The effects of knowledge spillovers on originating firms." In: *Academy of Management Journal* 53 (2), pp. 371–389.

Zirulia, Lorenzo and Nicola Lacetera (2010). "Knowledge Spillovers, Competition, and Taste for Science in a model or R&D incentive provision." In: *presented at the DRUID Conference*.