Université de Strasbourg

École Doctorale Augustin Cournot

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE AUGUSTIN COURNOT

Bureau d'Économie Théorique et Appliquée, UMR 7522

THÈSE présentée par :

Jia LIU

soutenue le : **21 Septembre 2022**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences de gestion - Innovation et entrepreneuriat

Knowledge Flow Within China and its Economic Consequences

THÈSE dirigée par :

M PÉNIN Julien	Professeur, Université de Strasbourg
M RONDÉ Patrick	Professeur, Université de Strasbourg

RAPPORTEURS :

Mme AUTANT-BERNARD Corinne	Professeure, Université de Saint Etienne
Mme LEVY Rachel	Maîtresse de Conférences, ENSFEA de Toulouse

AUTRES MEMBRES DU JURY :

M DEPRET Marc-Hubert	Maître de conférences, Université de Poitiers
M MULLER Paul	Maître de conférences, Université de Lorraine

ACKNOWLEDGEMENT

First and foremost, I have to thank my two supervisors, Professors Julien Pénin and Patrick Rondé. Without their assistance and dedicated involvement in every step throughout the process, this thesis would have never been accomplished. Their immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. I would like to thank you very much for your support and understanding over these past four years.

I would like to thank my colleagues at the Bureau d' Economie Theorique et Appliquee for their advice during my PhD. Special thanks to Patrick Llerena, Stefano Bianchini, Meixing Dai, and many others.

I thank Professors Corinne Autant-Bernard, Rachel Levy, Marc-Hubert Depret, Paul Muller, who do me the honour of accepting to be members of this jury.

I would like to express my gratitude to Jérôme Vicente and Stefano Bianchini for providing valuable suggestions that helped guide this entire project.

I would like to express my sincere gratitude to Université de Strasbourg for providing financial support without which this research would not have been possible.

Finally, my appreciation also goes out to my family and friends for their encouragement and support all through my studies. Their belief in me has kept my spirits and motivation high during this process.

Table of Contents

ACKNOWLEDGEMENT 2

Table of Contents 3

List of Table 5

List of Figures 6

Resume en Francais 7

General Introduction 16

Chapter I: Knowledge Spillovers Theory 22

I.1. Theory on knowledge spillovers 23

I.1.1. The concept of knowledge and knowledge spillovers and their classification 23

I.1.2. Knowledge Spillovers Mechanisms 30

I.1.3. Measurement of knowledge spillovers 47

I.2. Knowledge spillovers and innovation 56

I.2.1. Theory of technology, innovation and economic growth 56

I.2.2. Endogenous growth theory 59

Chapter II: Knowledge Flows Within Chinese Administrative Provinces: A Patent Citation Analysis 63

II.1. Introduction 64

II.2. Related literature on the determinants of knowledge flows 66

II.3. Empirical design 72

II.3.1. Data collection 72

II.3.2. The public versus private research structure of Chinese Provinces 74

II.3.3. Econometric treatment 78

II.4. Empirical results and discussion 83

II.5. Conclusion 87

Chapter Ⅲ: Knowledge Spillovers Through Co-patenting Among Inventors in China 99

Ⅲ.1. Introduction 100

Ⅲ.2. R&D cooperation theory 101

Ⅲ.3. Data and Method 103

Ⅲ.3.1 Data 103

Ⅲ.3.2 Model and variables 106

Ⅲ.4. Empirical Results 107

Ⅲ.5. Conclusion 109

Chapter VI: Knowledge Spillovers of Regional Innovation: The Cases Of China 119

VI.1. Introduction 120

VI.2. Knowledge Production Function 121

VI.3. Data and empirical methodology 122

VI.3.1. Data and variables 122

VI.3.2. Econometric approach 127

VI.3.3. The model 128

VI.4. Empirical results 130

VI.5. Conclusion 138

General Conclusion 140

References 144

Résumé 170

List of Table

Table 1 Factor analysis. 76

Table 2 Description of variables included in our model. 81

Table 3 Descriptive statistics. 82

Table 4 Regression results (1). 84

Table 5 Regression results (2). 85

Table 6 Number of inventors per patent. 112

Table 7 Descriptive statistics of dependent and independent variables（Obs：93,926） 113

Table 8 Correlation matrix of independent variables. 114

Table 9 Co-inventor dummy, 1995-2019 (patent level) 115

Table 10 Inter-provincial co-inventor dummy, 1995-2019 (patent level) 116

Table 11 Inventor count, 1995-2019 (patent level) 117

Table 12 Co-inventor max distance, 1995-2019 (patent level) 118

Table 13 Descriptive statistics of dependent and independent variables（Obs：434） 126

Table 14 Correlation matrix of independent variables. 126

Table 15 Moran's I for regional patents granted (2006–2019). 130

Table 16 Determinants of innovation based on the Spatial Durbin Model with spatial and time fixed-effects (2006-2019). 134

Table 17 Direct, Indirect effect, Total Effect Estimates of SDM (Data from USPTO) 135

Table 18 Direct, Indirect effect, Total Effect Estimates of SDM (Data from CNIPA) 136

Table 19 Determinants of innovation based on the two-way fixed effects model (2006-2019).137

List of Figures

Figure 1 . Factor 1 and Factor 2 scores of 31 provinces (The meaning of the abbreviation is in Appendix I). 77

Figure 2 . Comprehensive scores and ranking of regional innovation (following the factor analysis) 90

Figure 3 . Distribution of citations by technological field (IPC). 94

Figure 4 . Distribution of patents by co-inventor or not in the period of 1995-2019 111

Figure 5 . Distribution of patents by Intra-provincial collaboration and Inter-provincial Collaboration in the period of 1995-2019 111

Resume en Francais

Les Flux de Connaissances en Chine et Leurs Conséquences Économique

Cette thèse analyse les flux de connaissances par le canal des citations de brevets et de la collaboration des inventeurs sur les brevets dans le contexte de la Chine, puis estime leur contribution aux activités innovantes régionales.

Le premier chapitre donne un aperçu de la littérature existante sur la théorie des externalités de connaissances. Tout d'abord, nous expliquons la définition et la classification des externalités de connaissances. Le concept d'externalités de connaissances a été proposé dans les années 1960. Dougall (1960) a souligné que les externalités de connaissances sont un phénomène important des Investissements Directs Etrangers (IDE) lorsqu'on étudie les avantages sociaux des pays d'accueil qui bénéficient des IDE, et que les multinationales peuvent générer des externalités de connaissances et ainsi promouvoir la croissance de la productivité des entreprises nationales.

Arrow (1962) a expliqué le processus d'accumulation des connaissances et ses implications économiques. Il a indiqué que la connaissance a les caractéristiques de biens publics, et que la connaissance créée par une entreprise à travers des activités de R&D peut être facilement acquise par d'autres entreprises. Cependant, les innovateurs ne peuvent pas obtenir de compensation sous quelque forme que ce soit. Cette situation constitue ainsi une externalité de connaissances. Par conséquent, certains fabricants améliorent l'efficacité de leur production en investissant pour créer des connaissances, tandis que d'autres fabricants peuvent améliorer l'efficacité de leur production par l'imitation et l'apprentissage.

Romer (1986) a indiqué que le caractère partiellement excluable et non rival de la technologie est la principale raison du transfert de connaissances. On peut dire que Arrow et Romer ont apporté une contribution révolutionnaire au concept de transfert de connaissances.

La cause fondamentale des retombées de la connaissance est que la connaissance est un bien public. Une fois que la connaissance est produite, elle est essentiellement la richesse commune de toute la société. Les retombées du savoir peuvent être générées par l'investissement direct, le transfert de technologie, la coopération, la mobilité de la main-d'œuvre, etc., de manière consciente ou inconsciente, et peuvent être commerciales ou non commerciales (Grossman et Helpman 1991).

Dans un second temps, le chapitre 1 discute de cinq mécanismes de débordement des connaissances, notamment l'investissement direct étranger, le commerce, la coopération en matière de R&D, l'esprit d'entreprise et la mobilité de la main-d'œuvre. Ensuite, nous abordons le problème de la mesure des retombées du savoir à travers la fonction de production du savoir et l'approche par les citations. Enfin, d'un point de vue économique, nous discutons de la relation entre les flux de connaissances et l'innovation.

Le chapitre 2 analyse les déterminants des flux de connaissances dans le contexte de la Chine. Ce travail est l'une des premières tentatives d'application de l'analyse économique des retombées des connaissances en Chine, et il explore empiriquement les déterminants de la citation des brevets en Chine au niveau régional (au niveau des provinces chinoises). La plupart des recherches empiriques sur les retombées de la connaissance se concentrent sur les pays développés et le nombre d'études sur la Chine, bien qu'en augmentation, reste modeste. En outre, la plupart des études qui s'intéressent aux flux de connaissances en Chine se concentrent sur le niveau de l'entreprise (Xiang et al., 2013 ; Hansen et Hansen, 2020) ou tentent d'estimer l'impact des flux de connaissances provenant de l'étranger (Qiu et al., 2017). À notre connaissance, très peu d'études ont jusqu'à présent tenté de comprendre les déterminants des transferts de connaissances interrégionaux en Chine. Celles qui l'ont fait ont constaté que les collaborations intrarégionales et internationales sont les principaux canaux d'échange de connaissances en Chine, tandis que l'échange de connaissances interrégional est relativement faible (Gao et al., 2011). Il pourrait donc être important de comprendre les déterminants des flux de connaissances interprovinciaux en Chine afin d'élaborer des politiques susceptibles de débloquer les obstacles existants.

Se concentrer sur les flux de connaissances entre les provinces chinoises permet d'examiner des questions de recherche classiques telles que l'effet de la distance géographique ou de la proximité technologique entre deux provinces. Mais, plus originellement, cela permet également d'explorer dans quelle mesure les caractéristiques des provinces affectent les flux de connaissances. En particulier, dans ce chapitre, nous nous concentrons sur l'effet de la structure de recherche de chaque province chinoise.

Nous nous appuyons sur les citations de brevets pour mesurer et capturer les externalités de connaissances (Jaffe et de Rassenfosse, 2017). Depuis plusieurs décennies, les études économiques utilisent les statistiques sur les brevets pour mesurer l'activité innovante. Depuis les années 1990, les citations de brevets ont également été largement utilisées pour mesurer les débordements de connaissances (Jaffe et al., 1993). Les citations de brevets restent toutefois une mesure imparfaite des externalités de connaissance. Premièrement, parce qu'elles ne reflètent que les connaissances codifiées dans le document de brevet et négligent les connaissances contenues dans d'autres formes. Ensuite, parce que les examinateurs de brevets ajoutent souvent des citations au cours du processus d'examen, induisant ainsi un biais et du bruit dans la mesure des flux de connaissances (Jaffe et al., 2000 ; Alcacer et Gittelman, 2006 ; Alcacer et al., 2009 ; Lampe, 2012 ; Roach et Cohen, 2013 ; Moser et al., 2018 ; Corsino et al., 2019). Malgré cette importante mise en garde, pour Jaffe et de Rassenfosse (2019), il existe aujourd'hui un large consensus sur le fait que les citations de brevets restent un outil essentiel pour mesurer les caractéristiques des inventions (leur impact, leur valeur, leur généralité, leur originalité, etc.) et les flux de connaissances entre les acteurs économiques.

Dans la lignée de cette littérature, nous utilisons des données sur la localisation des inventeurs de brevets pour identifier la localisation géographique des brevets cités et citants. Chaque citation entre deux brevets est donc associée à deux des 31 provinces administratives chinoises (la province citante et la province citée). En nous appuyant sur les données de brevets de l'US Patent and Trademark Office sur la période 1995-2019, nous sommes en mesure d'identifier 27 118 paires de citations de brevets.

Nos résultats économétriques montrent que, comme prévu, la distance géographique et technologique entre les Provinces est négativement corrélée aux citations de brevets. En outre, nous constatons que les provinces administratives chinoises qui montrent une plus grande intensité en matière de recherche publique et privée reçoivent plus de citations de brevets et citent plus d'autres brevets dans d'autres provinces. Cet effet est particulièrement significatif lorsque les deux régions sont spécialisées dans la recherche privée, ce qui suggère que deux provinces chinoises fortement orientées vers la recherche privée sont plus susceptibles de générer des flux entrants et sortants de connaissances l'une avec l'autre. En revanche, cet effet n'est pas significatif lorsque les deux provinces sont spécialisées dans la recherche publique. Enfin, et en contradiction avec les recherches antérieures (Zucker et al., 1998 ; Breschi et Lissoni, 2003 ; Morrison, 2008 ; Frenken et al., 2010 ; Ott et Rondé, 2019), nous constatons que le contrôle de la proximité sociale entre les Provinces n'a pas d'impact sur l'effet de la proximité géographique et technologique. Ces résultats, s'ils sont confirmés par d'autres études, pourraient avoir des implications politiques importantes quant à l'influence de la structure de recherche des provinces chinoises sur la circulation des connaissances en Chine.

L'objectif du chapitre 3 est de combler les lacunes des recherches existantes en examinant dans quelle mesure les caractéristiques institutionnelles expliquent l'engagement des inventeurs dans la production collaborative de connaissances, et quels sont les modèles de collaboration entre organisations qui sont plus susceptibles de collaborer entre régions. L'importance des collaborations en matière de R&D pour l'innovation est largement reconnue depuis quelques années. La collaboration en matière de R&D contribue à l'innovation par le biais du partage des ressources et de l'expertise (Meli, 2000 ; Beaver, 2001 ; Sonnenwald, 2007), de l'échange d'idées (Melin, 2000 ; Birnholtz, 2007), de l'apprentissage de nouvelles compétences (Heinze et Kuhlmann, 2008), de la mise en commun de l'expertise pour des problèmes complexes (Sonnenwald, 2007) et de la facilitation des flux de connaissances (Singh, 2005 ; Montobbio et Sterzi, 2011). Les recherches existantes ont exploré la coopération en matière de R&D dans divers contextes. La collaboration en matière de R&D est principalement axée sur les aspects suivants : l'explication de la croissance (Monjon et Waelbroeck, 2003 ; Loof et Heshmati, 2002 ; Belderbos et al., 2015), la mesure de la collaboration en matière de R&D (Cronin et al, 2003 ; Savanur et Srikanth, 2009), les facteurs qui influencent les collaborations en R&D (Amabile et al., 2001 ; Birnholtz, 2007 ; Stokols et al., 2008 ; Bammer, 2008 ; Rigby, 2009), et les raisons de la collaboration (Beaver, 2001; Sonnenwald, 2007 ; Beaver, 2001). La collaboration avec divers partenaires de R&D permet aux entreprises de rechercher différents types de connaissances pour l'innovation. Des études empiriques récentes ont exploré l'influence de la coopération en R&D avec différents types de partenaires de collaboration, notamment les partenaires verticaux, les concurrents et les partenaires institutionnels (Aschoff et Schmidt 2008 ; Belderbos, 2015 ; Franco et Gussoni, 2014).

Les études existantes portent principalement sur la collaboration en matière de R&D via la publication d'articles cosignés (par exemple, Wagner 2005 ; Heinze et Bauer 2007 ; Mattsson et al. 2008) et parfois aussi sur l'enquête communautaire sur l'innovation. Cependant, la littérature spécifique à la collaboration des inventeurs en matière de brevetage est plus rare. Or, les brevets sont un élément important de la production inventive (Griliches, 2007). La plupart des recherches existantes sur le brevetage conjoint portent sur le réseau social connecté par les inventeurs. Les principales différences entre cette étude et les études précédentes sont les suivantes : Premièrement, nous utilisons le brevet comme unité d'analyse, contrairement à la plupart des études précédentes qui utilisent des données agrégées pour étudier la coopération des inventeurs au niveau de la ville, de la région. Deuxièmement, nous considérons différents types de collaboration ensemble dans un seul modèle.

Cette analyse se fonde sur les données relatives aux brevets déposés par les inventeurs chinois à l'USPTO entre 1995 et 2019 pour explorer les facteurs qui influencent les flux de connaissances par le biais de la collaboration des inventeurs en Chine en utilisant des modèles logit. Les principales conclusions de cette étude sont les suivantes : Premièrement, cette étude montre que les universités sont plus susceptibles de collaborer dans le domaine des brevets, et que les industries préfèrent les brevets de manière indépendante. Deuxièmement, les inventeurs affiliés à diverses universités ont une influence positive et statistiquement significative sur la probabilité d'une collaboration interrégionale en matière de R&D. Troisièmement, les inventeurs d'Université-Industrie et d'Inter-industries ont un impact négatif significatif sur la probabilité de collaboration interrégionale en matière de R&D. En d'autres termes, la coopération interindustrie et université-industrie est plus susceptible de se produire à l'intérieur des frontières provinciales. Cette constatation a des implications importantes en matière de politique économique.

Dans la littérature existante sur la diffusion régionale de la technologie, les principaux déterminants de l'innovation sont l'apport en R&D, le capital humain et les divers canaux qui facilitent les transferts de connaissances. Les études sur les retombées technologiques soulignent également que celles-ci sont, dans une large mesure, localisées géographiquement (Jaffe, 1989 ; Krugman, 1991). L'explication la plus fréquemment avancée pour la localisation régionale des connaissances est la nature tacite des connaissances, qui sont acquises par des contacts interpersonnels directs (Anselin et al., 2000). Étant donné que les retombées de la connaissance ne sont pas directement observables, il est difficile de trouver des preuves systématiques de l'étendue et de l'importance de ces impacts.

En prolongement des chapitres 2 et 3, le Chapitre 4 étudie comment les activités de R&D dans d'autres domaines influencent les performances d'innovation des provinces par le biais de divers canaux de diffusion des connaissances. L'objectif du chapitre 4 est d'apporter un éclairage supplémentaire sur cette question dans un contexte chinois. Ce travail est l'une des premières tentatives d'application de l'analyse économique spatiale des retombées des connaissances en R&D aux pays en développement. La contribution de cette étude à la littérature est d'examiner dans quelle mesure les retombées basées sur les citations de brevets et les connexions inter-personnelles contribuent à la production d'innovation à travers les régions.

Nous estimons les débordements de connaissances en Chine d'un point de vue temporel et spatial, au moyen d'une perspective spatiale en économétrie et d'un cadre de fonction de production de connaissances. L'approche de la fonction de production se concentre sur la production totale de connaissances en tant que fonction de l'apport en R&D et des retombées de connaissances. Plus précisément, cette recherche étudie le rôle de trois mécanismes de débordement des connaissances dans la détermination de la création de brevets en Chine : les débordements liés aux citations de brevets, les débordements liés aux contacts face à face basés sur la collaboration entre inventeurs, et l'apport en R&D des régions environnantes. Cette recherche estime un modèle de Durbin spatial, ne considérant pas seulement les estimations ponctuelles, mais incluant également les effets directs, indirects et totaux pour fournir des résultats plus précis.

Les conclusions de ce chapitre montrent que la diffusion des connaissances par le biais des citations de brevets et de la collaboration entre inventeurs, ainsi que l'apport en R&D, sont des ingrédients clés de la promotion des activités innovantes régionales. Ces résultats confirment l'importance des initiatives de R&D pour renforcer l'innovation en Chine et la nécessité d'encourager la collaboration pour améliorer l'efficacité de ces activités de renforcement de l'innovation. Par conséquent, les régions qui veulent rester compétitives au niveau international devraient investir dans les aspects de l'infrastructure de la connaissance de la région qui favorisent les retombées de la connaissance régionale. En outre, la présence de flux de connaissances spatiaux dans le processus d'innovation peut réduire les inégalités interrégionales en matière d'innovation.

General Introduction

Understanding the determinants of knowledge flows within the economy and the factors that might accelerate or hamper them is a critical task for at least three reasons. First, according to the new growth theory, knowledge flows are central to generate endogenous long-run growth (Romer, 1986; Acs et al., 2012). Second, the local versus global dimension of knowledge flows is considered an important explanation of the process of convergence or divergence between countries and regions and therefore significantly affects the possibility of catching up for developing countries (Abramovitz, 1986). Third, and linked to the former point, localized knowledge spillovers are viewed as a critical driver of economic geographic concentration (Feldman, 1994; Audretsch and Feldman, 1996; Audretsch and Feldman, 2004).

The important role of knowledge spillovers in economic growth is beyond question, and a large number of studies have explored the determinants of knowledge flows between economic actors. These flows are influenced by geographic proximity, technological and temporal proximity. Also, the observation of a significant national effect seems to indicate that institutional proximity plays a role (Boschma, 2004). However, there are two main challenges scholars confront. Firstly, to determine the mechanism or modes of communication that permit knowledge flow. The second challenge is how to identify and measure the knowledge spillovers (Audretsch and Feldman, 2004). Nadiri (1993) investigated the results of various countries, periods, and aggregation levels. Overall, the knowledge spillover effect is considered to be very significant and positive, but the estimated effect varied over a fairly wide range. One reason for the differences in effect lies in the methods used to measure spillovers between industries are quite different.

This thesis aims to fill the gap by focusing on the ‘geographic dimensions of knowledge spillovers’. According to these new ideas, the main contribution of this thesis to the literature is providing a comprehensive understanding of the underlying mechanism of knowledge spillover in China in combination with patent citation, inventor cooperation, and R&D input mechanisms. More precisely, the thesis first examines the influence of region characteristics, regional organizational research structures, and geographical boundaries on knowledge spillovers. Furthermore, we explore the extent to which institutional characteristics explain the inventor's engagement in collaborative innovative activities, and which organization collaboration patterns are more likely to collaborate across regions. Finally, we estimate the contribution of patent citation-related spillovers and face-to-face contact spillovers based on inventor collaboration on innovative activities.

This thesis aims to apply these theories to the case of the regions of China. In recent decades, China's innovation capability has improved significantly, and the trend does not show any sign of deceleration. In 2017, Chinese R&D intensity was equal to 2.12% of GDP and R&D total expenditure was equal to 280 billion US dollars, accounting for 20% of total world R&D expenditure. The total number of R&D personnel amounted to 3.759 million, accounting for 31.1% of the total world R&D personnel. China is also the country with the most patent applications in the world. The trend is similar with scientific research: China is now ranking second for the number of scientific publications authored by researchers working in China (Ministry of Science and Technology, 2017).

However, economic infrastructure and economic development in China are unevenly distributed across Provinces. Due to historical factors that persist or reinforce through time and differentiated regional development policy, significant differences can be observed in the distribution of regional innovation capabilities across Chinese Provinces. Therefore, identifying the main knowledge spillover mechanism and factors that influence knowledge spillover in China is critical to developing appropriate policies to encourage inter-regional knowledge flows, and maximize knowledge externalities. This thesis is structured as follows:

The first chapter gives an overview of existing literature on knowledge spillovers theory. First, we explain the definition and classification of knowledge spillovers. Second, we discuss five mechanisms of knowledge spillovers, including foreign direct investment, trade, R&D cooperation, entrepreneurship, and labor mobility. Then, we talk about the measurement problem of knowledge spillovers through the knowledge production function and paper trail approach. Finally, from an economic perspective, we discuss the relationship between knowledge spillovers and innovation.

The second chapter analyzes the determinants of knowledge flows in the context of China. Focusing on knowledge flows between Chinese Provinces enables the examination of standard research questions such as the effect of the geographical distance or technological proximity between two Provinces. But, and more originally, it also enables us to explore to which extent the characteristics of the Provinces affect knowledge flows. In particular, we focus on the effect of the research structure of each Chinese Province. We use factor analysis to measure the private versus public research intensity of each Chinese administrative province, thus introducing a distinction between Provinces that are more oriented toward public versus private research. This is in line with research that goes beyond the geographic or technological effect and looks at the institutional dimension in order to explain knowledge flows.

The results show that Chinese administrative provinces that show bigger intensity with regard to public and to private research both receive more patent citations and cite more other patents in other Provinces. This effect is particularly significant when the two regions are specialized in private research, thus suggesting that two Chinese Province strongly oriented toward private research are more likely to generate knowledge inflows and outflows one with each other. On the other hand, this effect is not significant when the two Provinces are specialized in public research. Finally, we find that controlling for social proximity between Provinces does not impact the effect of geographical and technological proximity. These results, if confirmed by further studies, might have significant policy implications as to the influence of Chinese Provinces' research structure on knowledge circulation within China.

The third chapter aims to answer two questions. The first set of questions aimed to analyze what factors affect the probability of China's inventors cooperating. The second is to explore which factor has an impact on cross-region knowledge flow through patent collaborations. More specifically, this chapter explores the extent to which institutional characteristics explain the inventor's engagement in collaborative knowledge production, and which organization collaboration patterns are more likely to collaborate across regions. This Chapter is based on the patent data applied by China's inventors in the USPTO between 1995 and 2019 to explore the factors that influence the knowledge flows through inventor collaboration in China using logit models.

The major findings of this chapter are as follows: First, universities are more likely to collaborate in patenting, and industries prefer patents independently. Second, inventors affiliated with various universities have a positive and statistically significant influence on the likelihood of interregional R&D collaboration. Third, the inventors from University-Industry and Inter-industries have a significantly negative impact on the probability of interregional R&D collaboration. In other words, inter-industry and UI cooperation are more likely to occur within provincial boundaries.

The fourth chapter investigates how R&D activities in other areas influence the innovation performance of Provinces through various knowledge spillovers channels. We estimate knowledge spillovers in China from a temporal and spatial perspective, by means of a spatial perspective in econometrics and knowledge production function framework. The production function approach focuses on the total output of knowledge generation as a function of R&D input and knowledge spillovers. More specifically, this research investigates the role of three mechanisms of knowledge spillovers play in determining the creation of patents in China: patent citation-related spillovers, face-to-face contact spillovers based on inventor collaboration, and R&D input from surrounding regions. This research estimates a spatial Durbin model, not only considering point estimates, but also including the direct, indirect, and total effects to provide more accurate results.

The finding of this chapter shows that knowledge spillovers through patent citations and inventor collaboration, and R&D input are key ingredients in promoting regional innovative activities. These findings confirm the importance of R&D initiatives for enhancing innovation in China and the necessity of encouraging collaboration to improve the effectiveness of such innovation-enhancing activities. Therefore, regions that want to stay competitive internationally should invest in those aspects of the region's knowledge infrastructure that promote regional knowledge spillovers. Moreover, the presence of spatial knowledge flows in the process of innovation can reduce interregional innovation inequalities.

Chapter I: Knowledge Spillovers Theory

I.1. Theory on knowledge spillovers

In this section, we introduce the concept of knowledge spillovers and the classification of knowledge spillovers, the knowledge spillovers mechanisms, and present a description of the measurement of knowledge spillovers.

I.1.1. The concept of knowledge and knowledge spillovers and their classification

I.1.1.1. The concept of knowledge and knowledge spillovers

Marshall(1890) first proposed the concept of "spillovers", and he equated the concept of spillovers with externality. He believed that in economic activities, the consumption of any scarce resources depends on the proportion of supply and demand. The root of economic inefficiency lies in external diseconomy.

In analyzing the development of localized industry, Marshall wrote "The mysteries of the trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously. Good work is rightly appreciated, inventions and improvements in machinery, processes, and the general organization of the business, have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas."

Pigou(1920) helped refine and develop Marshall's ideas, and argued that because of the existence of the divergence between social marginal net product and private marginal net product, the market forces could not be relied upon for optimal resource allocation, and government subsidization or taxes would be required to use as a tool for correcting inefficiencies in the allocation of resources in a competitive economy. And the external economy and external diseconomy are the positive and negative effects of spillovers respectively. Baumol (1952) proposed in the book "Welfare economics and the theory of the state" that the behavior of a certain manufacturer would affect other manufacturers in the same industry, and price changes could not compensate for this effect, thus creating a spillover effect. Bator tried to broaden the definition of externalities to include all types of market failures, he stressed:

"In its modern version, the notion of external economies–external economies proper that is: Viner's technological variety–belongs to a more general doctrine of 'direct interaction'. Such interaction consists of inter-dependences that are external to the price system, hence unaccounted for by market valuations. Analytically, it implies the non-independence of various preference and production functions." (1958, 358)

Stiglitz (1997) defined spillover as "excess costs and benefits not included in market transactions", which refers to a phenomenon that occurs when an individual or a manufacturer does not bear the full cost of its actions or does not enjoy all of its benefits.

"R&D spillovers refer to the involuntary leakage, as well as, the voluntary exchange of useful technological information." (Steurs, 1994: p. 2) The concept of knowledge spillover was proposed in the 1960s, Dougall (1960) firstly clarified the concept of knowledge spillovers. He pointed out that knowledge spillover is an important phenomenon of FDI when studying the social benefits of host countries accepting foreign direct investment, and MNCs can generate knowledge spillovers due to the externalities of the economy and promote the growth of the productivity of domestic firms. Arrow (1962) firstly explained the process of knowledge accumulation and its economic implications. He indicated that knowledge has the characteristics of public goods, and the knowledge created by a company through R&D activities can be easily acquired by other companies, however, innovators cannot obtain compensation in any form. This situation is knowledge spillover. Therefore, some manufacturers improve their production efficiency by investing to create knowledge, while other manufacturers can improve production efficiency through imitation and learning. Romer (1986) indicated that the partially excludable, non-rival character of technology is the main reason for the knowledge spillover. It can be said that Arrow and Romer made a groundbreaking contribution to the concept of knowledge spillover.

Kokko (1994) defined knowledge spillover in the foreign investment context as “the knowledge possessed by a foreign company is acquired by a local enterprise without the formal transfer.” The spillover effect comes from two aspects: the first comes from demonstration, imitation, and dissemination, and the second comes from competition. Griliches (1992) defined knowledge spillovers as “working on similar things and hence benefiting much from each other’s research.” Jaffe and Trajtenberg (1996) proposed that knowledge spillovers occur because knowledge created by one firm typically is not contained in that firm; therefore, it creates value for other firms and other firms’ customers. As Branstetter (1998) argued, knowledge spillovers are “an enterprise that can derive economic benefits from the R&D activities of another firm without sharing the research costs of the other.”

Caniëls (2000) emphasized the intellectual gains by exchange of information with a lack of direct compensation or at least less compensation than the value of the knowledge to the producer. Knowledge spillovers can be defined as the amount of knowledge that can’t be appropriated by the economic agent who created it (Greunz 2003). Fallah and Ibrahim (2004) distinguished between “knowledge spillovers” and “knowledge transfer”, and proposed that spillovers are the unintentional transmission of knowledge to others beyond the intended boundary. At every possible interaction, there is a potential for knowledge exchange. If knowledge is exchanged with the intended people or organizations, it is “knowledge transfer”, any knowledge that is exchanged outside the intended boundary is spillover. The unintended "use" of exchanged knowledge is called "Knowledge Externality".

Similarly, Grossman and Helpman (1991) believed that the root cause of knowledge spillovers is that knowledge is a public good. Once knowledge is produced, it is essentially the common wealth of the whole society. For the reason of interests and competition, some knowledge has a strong monopoly. Knowledge spillovers can be generated through direct investment, technology transfer, cooperation, labor mobility, etc., either consciously or unconsciously, and can be commercial or non-commercial.

I.1.1.2. The classification of knowledge spillovers

According to whether requiring material carriers, knowledge spillovers can be divided into two types: pecuniary spillovers, which produce their effect through market dealings, and relate to the purchase of equipment, goods, and services; and pure knowledge spillovers, which are independent of any market mechanism (Griliches, 1979; Verspagen, 1991). Pecuniary knowledge spillovers occur when a new or improved product is sold, but the innovating industry can't fully appropriate the increased quality of their products, then part of the productivity gains made by the innovating industry, finally belongs to downstream industries (shifting rents from innovators to users). Indeed, pecuniary spillovers occur in a not perfectly competitive market.

Pure knowledge spillovers are pure technological externalities, which are not embodied in a particular service or product, though they might be conveyed by a printed article or a news release. It has the classic aspect of a non-rivalrous good and it is usually difficult to appropriate more than a tiny fraction of its social returns (Griliches, 1991).

The scope of knowledge spillover can be divided into (1) Knowledge spillovers within the organization; knowledge spillovers between organizations; (2) knowledge spillovers between individuals; knowledge spillovers between individuals and organizations; (3) knowledge spillovers between industrial districts; knowledge spillovers within the industrial district; and (4) domestic knowledge spillovers; international knowledge spillovers. The externality is the most fundamental feature of knowledge spillovers. The term 'externalities' refers to economies of scale external to the firm. According to the external characteristics of knowledge, the knowledge spillovers are identified in three forms: MAR externality; Porter externality; and Jacobs externality.

MAR externality is developed by Marshall (1890), Arrow (1962), and Romer (1986), formalized by Glaeser et al. (1992) as MAR externality, which is the knowledge spillovers between enterprises in the same industry. As Glaeser (1992) emphasized, "knowledge accumulated by one firm tends to help other firms' technology without appropriate compensation". MAR externality can increase the industry's productivity in the agglomerated region, mainly through the imitation of innovative products, reverse engineering, business interactions, and mobility of skilled labor. MAR externality emphasizes that high local specialization and local monopoly can encourage economic growth, because the lack of property rights to new ideas will discourage innovators to invest in externality-generating activities, and monopoly permits the externality to be internalized by the innovator. At the same time, the competitive market and diversity are bad for growth.

Porter externality is proposed by Porter(1990), and it is similar to MAR externality. Porter externality indicates that knowledge spillovers occur most effectively between enterprises in the specific industry geographically concentrated. Compared with MAR externality, Porter externality believes that the competitive market, rather than a monopoly market structure, is more likely to promote the exchange of information, knowledge, ideas, and innovation. There is a lack of market pressure in the monopoly market, and entrepreneurs are reluctant to invest in risk innovation, thus, externality can be maximized in an industry-specific, competitive market structure. As Glaeser et al.(1992) stressed, “firms that do not advance technologically are bankrupted by their innovating competitors, although such competition reduces the returns to the innovators”. So the firms have to innovate in the competitive environment.

On the other hand, Jacobs externality arises from the idea and analysis of Jacobs (1969), who investigated the economic history of various cities. In contrast to the MAR externality and Port externality which are generated by enterprises in the same industry, Jacobs externality is usually generated when spillovers occur between enterprises in different industries. Industry diversity within a geographic region, rather than geographical specialization, is the major engine for city prosperity and economic growth, and the exchange of complementary knowledge across different industries can facilitate innovation. Jacobs externality believes that competition is beneficial to knowledge spillovers, and regional competition accelerates the industrialization of new technologies. Jacobs (1969) argued that the most important source of knowledge spillovers is external to the industry in which the firm operates and that cities are the source of considerable innovation because the diversity of these knowledge sources is greatest in cities. According to Jacobs, it is the exchange of complementary knowledge across diverse firms and economic agents which yield a greater return on new economic knowledge. She develops a theory that emphasizes that the variety of industries within a geographic region promotes knowledge externalities and ultimately innovative activity and economic growth.

In general, the commonality of the three externalities lies in the industry's geographical agglomeration promotes knowledge spillover and regional innovation. However, there also exist some differences. Firstly, the sources of externality are different. MAR externality and Port externality come from the same industry, while Jocbos externalities come from different industries. Secondly, the impact of regional competition or monopoly on externality varies widely. MAR externality regards competition as an obstruction to innovation. In contrast, Port externality and Jacobs externality regard competition as a positive factor for innovation. Lastly, The economies of scale of knowledge spillovers come from the MAR and Port externalities of knowledge. However, the economies of scope of knowledge spillovers stem from the Jacobs externalities of knowledge.

There are some empirical studies trying to test for the existence of the three forms of knowledge spillovers. Glaeser et al. (1992) selected 170 metropolitan areas in the United States to conduct research on the six largest industries between 1956 and 1987. The result is consistent with the Jacobs externality theory that local competition and diversity, but not regional specialization, encourage employment growth in industries. Henderson (1995) selected data from 1970-1987 in 224 metropolitan areas in the United States, involving five traditional industries and three high-tech industries., and the conclusions show that MAR externality has been found in some traditional industries, and both MAR and Jacobs externality in high-tech industries played a significant role in knowledge spillovers. In the Italian case, Paci and Usai (1999) concluded that the diversity and specialization externality both have a positive effect on innovative activities, and Jacobs externality is more powerful in the high-tech sector and in metropolitan areas. Audretsch and Feldman (1998) tested whether the specialization or diversity of economic activity is more conducive to knowledge spillovers, and the results indicate that diversified across complementary industries tends to have more innovative activities. In the Dutch context, Van der Panne (2004) supported the opinion of

MAR externality theory that regional specialization tends to encourage regional innovative activities in that industry.

I.1.2. Knowledge Spillovers Mechanisms

I.1.2.1. Foreign direct investment

Since the 1980s, Foreign Direct Investment (FDI) has become a major form of technology transfer from developed to developing countries. The FDI spillover effect refers to the fact that a multinational corporation's establishment of a subsidiary in the host country leads to an increase in local technology or productivity, while Multinational Corporations (MNCs) can't obtain all the benefits. MNCs have been linked to some advantages such as superior technologies, patents, trade secrets, brand names, management techniques, and marketing strategies (Dunning, 1993). Consequently, many countries offer various preferential policies such as tax exemptions, subsidies, and low tax rates to attract foreign investors.

Dunning (1998) points out that FDI can promote knowledge diffusion for the reasons below：(1) The parent company's knowledge advantage is the basis for the subsidiary to gain a competitive advantage in the host country, and the parent company is often willing to share international advanced knowledge with the subsidiary; (2) The subsidiary can use the various network relationships that the parent company has formed in the world to gain more external knowledge sources. (3) Although the subsidiary is not obliged to spread knowledge to the host country, it will use knowledge transfer of resources and involuntary knowledge spillover demonstration effects and human capital flows make knowledge spillovers to the host country.

The pioneering study of the FDI productivity spillovers is from Caves (1974), who examined FDI in manufacturing industries in Australia and Canada. The result shows that domestic Australian firms that compete in industries with a high FDI presence will have higher productivity. The research by Globerman (1979) on Canadian manufacturing, Blomstrom and Persson (1983) on Mexican manufacturing, and research by Flores et al. (2000) on Portuguese manufacturing show that the spillover effect of FDI is obvious. However, the study evidence from Morocco (Haddad and Harrison, 1994), Uruguay (Kokko et al., 1996), Spain (Barrios and Strobl, 2002), and the UK (Haskel and Slaughter, 2002) have a converse conclusion that the spillover effect of FDI does not exist. Through a comparative analysis of 101 countries, Blomstrom et al. (1994) found that the spillover effects of FDI mainly occurred in middle-income developing countries, while in the poorest developing countries, no evidence was found to prove the existence of such spillover effects.

Knowledge spillovers occur from MNCs to host firms mainly in five ways: linkage effect, demonstration effect, training of employees, competition effect, and export. The linkage effect refers to domestic firms entering into partnerships with multinationals. Demonstration effects refer to domestic firms replicating or imitating the technologies and processes of foreign firms to achieve technological improvement; MNEs often invest in training local employees, and the training effect occurs if there is employee turnover between MNEs and domestic firms; However, it is important to emphasize the negative impacts may arise through this channel because the foreign firms can attract the skilled labor in domestic firms by offering higher wages (Sinani and Meyer, 2004). The competition effect is the changes induced by foreign investment among domestic competitor firms in the host economy. MNCs can eliminate monopoly to a certain extent, and the competition forces domestic firms to explore efficient uses of existing technologies or to develop new technologies. On the other hand, domestic firms that can’t compete with foreign firms are forced to withdraw their investment (Blomström and Kokko, 1998). With the data of manufacturing in Ireland to analyze the impact of FDI on the entry of domestic firms in host economies, Barrios et al. (2004) concluded that at the beginning the competition effects dominate, but as time goes by, the positive externality effects associated with MNEs will have positive effects on domestic firms' start-up. Blake et al. (2009) used the data of 998 manufacturing firms to identify the main FDI spillover channels in China and suggested that the export of MNEs is the most important channel. Moreover, employee mobility, demonstration, and competition effect generate positive spillovers to the local firms which have greater absorb capability. Orlic et al. (2018) stressed that the demonstration effect has a negative impact on the productivity of domestic firms.

Nevertheless, the conclusion that the effect of FDI on backward spillovers and horizontal and forward spillovers varies in the empirical studies and has been rather inconclusive. Many pieces of evidence indicate that spillovers are more likely vertical than horizontal spillovers (Blalock and Gertler, 2008; Javorcik, 2004; Newman et al., 2015). Havranek and Irsova (2011) applied a meta-analysis of data from 3626 estimates of vertical spillovers to indicate that the average spillover to suppliers (backward spillovers) is economically significant, while the spillover to buyers (forward spillovers) is statistically significant but small. Lin and Saggi (2004) investigated the effect of FDI on backward linkages, and the results suggested that MNCs have two opposite effects on the upstream industry: on the one hand, MNCs increase the demand for intermediate goods, which leads to increased output from local suppliers; On the other hand, MNCs indirectly reduce the demand for intermediate goods by crowding out local competitors to exit the final good market. Moreover, the study believed that the significance of the backward spillover effect depends on the technological gap between MNCs and domestic firms, the greater the technological gap, the smaller the backward spillover effect. Merlevede et al. (2014) analyzed the case in Romania and concluded that most MNCs increase the productivity of local suppliers a few years after entry, then the effect fades out after the foreign firm is present for more than 3 years, thus, the effect on the productivity of local suppliers is transient.

Using a sample of 11767 Hungarian firms, Békés et al. (2009) found that the entry of MNCs in the same industry boost competition and reduces the productivity of the least productive local firms, conversely, more productive local firms benefit more from horizontal spillovers. In the case of China, positive forward spillovers arise when domestic firms buy higher quality intermediate inputs or equipment from foreign firms. For 17 emerging market economies, Gorodnichenko et al. (2014) suggested that the spillover effect of backward linkages is always positive, the forward spillover effect is only positive to the old and service firms, while the horizontal spillover effect is insignificant but positive; and Halpern and Murakozy (2007) found positive vertical and negative horizontal FDI spillovers effect in Hungary. Applying the data of 10 transition countries, Damijan et al. (2013) showed that horizontal spillovers have become increasingly important in the past decade, and positive horizontal spillovers are equally distributed among firms of all sizes, meanwhile, negative horizontal spillovers are more likely to occur in smaller firms. Orlic et al. (2018) indicated that local manufacturing firms benefit from the backward spillovers in manufacturing sectors and forward spillover effects of FDI in services sectors.

Consequently, the FDI spillovers have a positive or negative impact on the productivity of domestic firms depending on whether the negative competition effect outweighs the positive effect of demonstration and imitation, the training of employees, and the linkages effect.

The literature on spillover effects shows that the impact of FDI on domestic firms depends on a variety of factors. Generally speaking, the characteristics of the domestic firm, MNCs and the host country environment are regarded as the most important factors, such as the absorptive capability of the domestic firms, as well as the technological capability, embeddedness, and autonomy of MNCs.

Using the data in Estonia in the period 1994-1999, Sinani and Meyer (2004) found that the FDI spillovers effect is quite large in Estonia, and labor and sales-intensive MNCs generate more spillovers than equity-intensive MNCs. On the other hand, the spillovers effects will be influenced by the characteristics of local firms, such as the size, the type of trade, and the equity structure. Hermes and Lensink (2003) found that having a developed financial system is linked to the positive FDI effect on economic growth. Estimating horizontal spillovers from FDI for 45 countries, Iršová and Havránek (2013) suggested that when the technology gap between domestic firms and foreign firms is too large, horizontal spillovers are small. On the other hand, a higher level of human capital in the host country is associated with larger spillovers. Javorcik (2004) found evidence that joint ventures are more likely to engage in local sourcing than wholly-owned subsidiaries, thus, leading to greater spillovers. Jordaan (2005) introduced the data of Mexican manufacturing industries to identify the determinants of FDI spillovers, and the results showed that agglomeration has a positive impact on FDI spillovers in Mexico. Thus, geographical proximity between MNCs and domestic firms can be regarded as a determinant of the FDI spillovers effect. Todo and Miyamoto (2006) indicated that the effect of FDI spillover varies greatly depending on the types of activities undertaken by MNCs, especially in R&D activities, which is an important source of spillover.

Domestic firms with more R&D investment will benefit more from the presence of MNEs (Blalock and Gertler, 2009). Javorcik, Saggi, and Spatareanu (2004) studied the backward linkages spillover effects of FDI and found that the share of MNEs purchasing intermediate goods from the local market may be positively affected by the distance between the host country's production plant and the foreign intermediate goods production plant. Lenaerts and Merlevede (2015) conducted a study of Romanian firms and found that only medium-sized foreign firms can generate significant backward linkage spillover effects. Large foreign firms have less embedding of host economies and are more likely to use their own supplier systems to import intermediate products. Smaller foreign firms lack the scale of delivering spillover effects to domestic firms.

I.1.2.2. Trade and Knowledge spillovers

Theoretical studies indicate that trade and economic openness play an important role in transferring technology between two countries. Through international trade, a country can import more diverse or higher-quality intermediates, manufactured goods, and capital equipment embodying foreign knowledge; imitate and learn more advanced production processes, organizational routines, and management experiences from abroad; stimulate local firms to pursue more innovative ideas and technologies; allocate effectively domestic resources and reduce duplication of R&D. Thus, trade can boost the innovation capability of a country and increase domestic production.

Furthermore, trade not only influences technological innovation through knowledge spillovers, but the transfer of traded goods (especially the import of capital goods and intermediate goods) directly increases production efficiency. The technology is embedded in mechanical equipment, and imported machinery can immediately increase productivity.

Grossman and Helpman (1991) firstly introduced the general equilibrium model to analyze the relationship between trade, growth, and technological changes in an open economy and explained the impact of intermediate product trade and final product trade on long-run growth. In the model of product variety, the number of intermediate product types available in the market and the share of labor employed in the production process stimulate the total factor productivity growth of the country. In their analytical framework, technology spillovers occur through the import of intermediate goods. If a country produces new intermediate products that differ from existing intermediate products or are better than existing intermediate products, when these intermediate goods are exported, the productivity of importing countries will be improved through the R&D effect and technology diffusion of their trading partners, thus promoting the technological improvement and growth of the importing countries. Moreover, the research shows that changes in the trade openness of a country, as measured by trade promotion or trade protection levels can also affect the growth rate.

In the empirical analysis, Coe and Helpman (1995) is the seminal paper to test the relationship between domestic R&D, foreign R&D, and domestic total factor productivity (TFP) growth. In order to measure the foreign knowledge stocks, A variable is constructed as a weighted sum of the cumulative R&D expenditures of the country's trading partners, where the weights are given by the bilateral import shares. Using data on 21 OECD plus Israel during the period between 1971 and 1990, this analysis supports the view that the international R&D spillovers are trade-related. The research implies that an increase of 1% of R&D expenditures in developed countries will increase output in developing countries by 0.6%; almost half of this is brought by the United States because it is a trading partner of many developing countries and its R&D expenditures was the largest; it proved that the R&D expenditures in developed countries bring great benefits to developing countries. The result proved that: Firstly, a country's total factor productivity depends not only on its own R&D capital stock but also depends on the R&D capital stocks of its trade partners; Secondly, the more the economy is open to international trade, the stronger the impact of foreign R&D

capital on domestic productivity. Using a sample of 77 developing countries and 22 industrial countries in the period 1971-1990, Helpman and Hoffmaister (1997) further presented empirical evidence that R&D investment in industrial countries' trade partners has a significant positive impact on the growth of total factor productivity in developing countries.

In other words, Falvey et al. (2002) used the data of five OECD countries and 52 developing countries to examine the presence of knowledge spillovers and the results suggested that the level of trade is important in facilitating knowledge spillovers from industrial countries to developing countries and whether regard knowledge as a private good or public good is very important for receiving countries. If knowledge is a public good, the knowledge spillovers effect of import trade is more significant. Moreover, an increasing 1 percent in the knowledge stock of the industrial countries increases growth in the developing countries by between 0.01 and 0.07 percent in the short-run. The analysis of Connolly (2003) proved that imitation and innovation in developing countries are positively affected by imports from developed countries and that the role of import channels is more important for developing countries than for developed countries. Yasar and Paul (2007) estimated the relationships between productivity and FDI, trade, and licensing in Turkish manufacturing firms, and the research indicates that productivity is most closely related to FDI, especially for large companies; followed by exports and license trade, and finally imports of intermediate goods. It can be found that the effect of imports on Turkey's productivity increase is small.

Using a sample of the OECD countries for over 135 years, Madsen (2007) indicated that trade-related spillovers have a significant impact on the total factor productivity growth, Similarly, trade-related spillovers contribute significantly to the total factor productivity convergence in the OECD countries. Loecker (2007) applied a matched sampling technique to examine whether companies engaged in exporting become more productive in Lovenian manufacturing firms. The analysis supports the view that export increases the productivity of export entrants significantly. Furthermore, the productivity gap between exporters and their domestic competitors increases further as time goes by. Moreover, firms exporting their goods to high-income countries gain higher productivity. Coe et al. (2009) reestimated their model in Coe (1995) and considered the criticisms of other authors, and their analysis still supports the view that the impact of domestic and foreign R&D capital stocks on TFP. In addition, they argued that institutional differences are important determinants of TFP and that they impact the degree of R&D spillovers. The effect of R&D spillovers is stronger with institutional factors like strong patent protection, a specific origin of the legal systems, the ease of doing business, and the high quality of tertiary education systems.

Applying the data of 24 advanced countries in the period 1971 to 2004, Fracasso and Marzetti (2015) provided evidence that trade positively affects international knowledge spillovers. In addition, the sheer size of bilateral trade flows seems to be related to larger knowledge spillovers. Moreover, particularly intense trade flows generate greater international knowledge spillovers. Keller (2010) constructed a model to examine the contribution of international trade and FDI to economic growth. The findings show that there exists geographic localization of R&D spillovers and there is evidence for technology R&D spillovers through international trade and the activity of MNCs. In addition, the technology spillovers channels of international trade and FDI are correlated. Applying a sample of 20 European countries between 1995 and 2010, Ali et al. (2016) found evidence for the effects of FDI and trade spillovers on domestic productivity. Additionally, there is a strong complementary relationship between R&D spillovers through the channels of trade and FDI.

I.1.2.3. R&D cooperation and Knowledge spillovers

R&D institutions in universities and R&D departments of firms are regarded as important sources of knowledge spillover by endogenous growth theory. The definition of collaboration is “the process through which two or more actors engage in a constructive management of differences in order to define common problems and develop joint solutions based on provisional agreements that may coexist with disagreement and dissent” (Hartley et al., 2013: 826). The R&D cooperation process can transfer knowledge, exchange resource, and organizational learning. Particularly, through the network established by formal or informal face-to-face interaction, entrepreneurs, university researchers, and firm R&D employees can exchange information and knowledge, and make knowledge spillovers occur. R&D Collaboration has several advantages, such as achieving economies of scale; joint financing of R&D, acquisition of additional resources, reducing uncertainty, and cost savings; (Camagni, 1993; Robertson and Langlois, 1995; Becker and Dietz, 2004).

As for the impact of collaboration on the innovative performance, horizontal spillovers (competitors) may improve or damage innovation performance significantly in line with the types of R&D cooperation, while vertical spillovers (suppliers and customers) increase the R&D performance slightly on all occasions. Nevertheless, once exist, the benefits of horizontal spillover effects are greater than vertical spillover effects. It implies that horizontal cooperation in R&D has a substantial effect (Atallah, 2002). Zhang and Tang (2017) used patent data from 39 companies in China to investigate how intra-firm collaboration influences innovation performance. The result shows that the relationship between the scope of collaboration and innovation performance is positively impacted by the technological heterogeneity of employees. Becker and Dietz (2004) studied the case in Germany and pointed out that R&D cooperation can increase the R&D input intensity of enterprises, and also can increase the chance to develop new products.

Innovation collaboration is related to the acquisition of knowledge that is not available within an enterprise (Bogers et al., 2018). Most studies hold the opinion that R&D collaborations complement rather than compensate for localized R&D spillovers. Investigating quantitative data of the firms in Sweden, Grillitsch and Trippl (2014) pointed out that firms, which is less accessible to local knowledge spillovers, prefer to collaborate more. Similarly, using the firm-level data in German manufacturing, Schmiedeberg (2008) found there are significant complementarities between internal R&D and R&D cooperation. Beers et al. (2008) explored the R&D collaboration in Finland and the Netherlands, and the research implies that the incoming R&D spillovers are a significant factor for R&D collaboration with domestic public R&D institutions. Investigating the joint publications and co-patents in EU regions, Hoekman et al. (2009) showed that there are elite structures between the capital regions and the regions of excellence. Also, the constraints to innovation collaboration do not rise with the geographical markets where this collaboration takes place (Audretsch and Belitski, 2022).

To explore the characteristics of firms that are more likely to engage in R&D collaboration, Veugelers and Cassiman (2005) used the manufacturing firm-level data of Belgian and concluded that the industry sector is a significant factor. Particularly, in science-based industry such as chemicals and pharmaceuticals, firms tend to engage in R&D collaboration activities. Large and R&D-intensive enterprises are more collaborative than others. De Faria et al. (2010) explored how the choice of cooperation partners affects the innovation ability and found that the joint partner's type, innovation intensity, absorptive capacity, spillover management, and innovation ability of industry all affect the contribution of partners to R&D collaboration activities. Focused on how institutional logics of R&D activity, such as the rule-based, normative, and cultural logics impact the collaborative process between universities and SMEs. Contrary to the conclusions of much literature, Bjerregaard (2010) holds the opinion that institutional logic can promote continuous knowledge exchange. In addition, many collaborating researchers have undergone an institutional convergence constituting a shared cultural space for knowledge exchange in the collaborative process of R&D activity. Fritsch and Lukas (2001) found that firms involved in R&D cooperation are more likely to have a relatively large size and a high proportion of R&D investment. Comparing the cooperative innovation activities in Austria and Finland, Dachs et al. (2008) recognized that sectoral affiliation plays an important role to determine whether to conduct R&D collaboration in Austrian firms, however, in Finland, the innovation strategy of firms is the key factor.

While collaboration within firms, research centers, and universities remains crucial, external networks are a key feature of innovation teams (Crescenzi, 2016). Nieto and Santamaría (2007) concluded that the innovation collaborative networks made up of various types of partners has the most positive influences on innovation performance. As an important channel of knowledge spillovers, formal R&D collaboration can occur across long distances through the networks instead of only localized spillovers, thus making knowledge spillovers at the national level or international level (Ponds et al., 2009; Hoekman et al., 2009).

Additionally, the R&D collaboration relationship between firms and universities is based on the firms' complementary strategies. Santoro et al. (2000) showed that the intensive the relationships between the UIC, the higher the advancement of both knowledge and new technologies; and if intensive relationships can be established at the beginning of the cooperation, it will contribute more to innovation performance than the intensive relationships which are gradually established in the later stage. Investigating the case of France firms, Monjon (2003) found that highly innovative firms seem to benefit the most from R&D collaborative activities with foreign universities. Exploring the function of technology-based enterprises in university/industry collaboration (UIC), Motohashi (2005) pointed out that the UIC has developed rapidly in small enterprises in the past five years, and the impact of UIC on the productivity improvement of small enterprises is far greater than that of large enterprises. Using the data of firms in Italy, Medda et al. (2006) concluded that R&D cooperation with other firms imposes a positive impact on productivity, whereas R&D cooperation with universities doesn't improve productivity. Lööf (2009) investigated a sample of 2071 firms in Sweden, and the results showed that R&D collaboration with universities imposes a positive impact on the innovation performance of large manufacturing firms, and there is no significant correlation between university collaboration and the innovation performance of average service firms. Guan et al. (2006) analyzed 950 firms in Beijing, and the evidence supports the view that the degree of novelty of industrial innovation is directly proportional to the degree of R&D cooperation between universities and firms. It also points out that the incentive effect of UIC cooperation on industrial innovation is still in an inefficient stage in China.

We furthermore discuss R&D cooperation between different institutions in chapter Ⅲ.

I.1.2.4. Entrepreneurship and Knowledge spillovers

Entrepreneurship refers to the startups and development of new enterprises and a method of diffusing and converting knowledge into societal utility. Acs et al. (2009) developed a Knowledge Spillover Theory of Entrepreneurship (KSTE), arguing that entrepreneurship can be viewed as an intermediary through which the knowledge generated by incumbent firms can flow to the start-ups where it is commercialized. In detail, opportunities arise when the incumbent firms invest in new knowledge but do not commercialize it, and entrepreneurship is a method to capture such opportunities and get more benefits from these new ideas.

Entrepreneurship may stimulate innovation, economic growth, employment, and competitiveness (Acs et al. 2012, González-Pernía et al., 2012).

As demonstrated in the study of innovative UK enterprises by Audretsch et al. (2021), external knowledge-based methods can serve as a useful and supplementary source of knowledge for innovative organizations, particularly start-ups. Entrepreneurs who create a business in an agglomeration can gain more tacit knowledge. Engaging entrepreneurs in the entrepreneurial process and interacting with other organizations, especially collaboration, can help generate knowledge spillover.

Endogenous growth theory argues that knowledge spillover occurs spontaneously in the production process, but the KSTE thinks knowledge spillover occurs as a result of entrepreneurs' intentional absorption and it is not spontaneous. The KSTE shifts the unit of analysis from the enterprise level to the individual level, such as engineers, scientists, and others. Entrepreneurship is related to an individual's entrepreneurial abilities, which involve innovative ideas, adventurous spirit, and cooperation abilities. Individuals with entrepreneurial talents are good at identifying entrepreneurial opportunities and gaining tacit knowledge in their knowledge environment. In addition, entrepreneurship activity is frequently quantified by the rates of new firm start-ups, self-employment, business ownership, and a combination of start-ups and turbulence.

In the empirical research, Zucker et al. (1998) explored the geographic distribution of star scientists and newly established U.S. biotechnology firms and showed that star scientists in universities can transfer their knowledge to new ventures by pursuing private commercial interests as entrepreneurs, and the star scientist plays an important role in the knowledge spillover process. Callejon and Segarra (1999) concluded that both entry and exit rates of start-ups influences positively on the growth of TFP in industries and across regions.

Through empirical tests, Feldman (2001) found that the entrepreneurial spirit of entrepreneurs has an important impact on the agglomeration of innovation activities in a certain area and knowledge spillover. Braunerhjelm et al. (2010) also provided evidence to support the view that entrepreneurs contribute to economic growth and knowledge spillover can occur through both incumbents and entrepreneurial activities. Furthermore, policies of stimulating entrepreneurship are important tools for promoting knowledge spillovers and economic growth. Acs et al. (2012) used the entrepreneurship data of 18 countries and emphasized that apart from R&D and human capital, entrepreneurial activities also contribute to knowledge spillovers and economic growth. Plummer and Acs (2014) found that there is a positive relationship between new knowledge and entrepreneurship, which is negatively moderated by localized competition. González-Pernía et al. (2012) studied the case in Spanish regions and the results indicated that the region with a higher capacity for creating new knowledge and entrepreneurial activities is positively associated with its level of competitiveness.

As for the spatial dimension, it is widely believed that entrepreneurial activity varies across regions. Audretsch and Fritsch (2002) pointed out that differences in the characteristics of regional location have an important impact on new firm start-ups. Audretsch and Lehmann (2005) found that the number of start-ups located around universities is positively associated with the knowledge production of universities. The evidence supports the view that the knowledge spillover theory of entrepreneurship holds for both the regions and industries. Gilbert et al. (2008) indicated that start-ups created within geographic cluster locations can acquire more knowledge and perform better in terms of growth and innovation. Through studying the case in Germany, Fritsch and Aamoucke (2013) found that start-ups of innovative industries benefit a lot from the regional universities, public research institutes, and private R&D sector, however, such a situation doesn’t exist in the non-innovative industries. This conclusion is consistent with previous research, by Bade and Nerlinger (2000), and Harhoff (1999). Saxenian and Hsu (2001) found evidence to support the view that immigrant entrepreneurs are regarded as middlemen for two-way spillovers of the idea, skill, and knowledge between Silicon Valley and their countries of origin in Asia.

I.1.2.5. Labor mobility and knowledge spillovers

Human capital refers to the knowledge and skills that are condensed on employees, and this knowledge and skills can, to some extent, increase social labor productivity, promote economic growth, and realize value-added in the process of production, which brings surplus value. This form of non-material capital expressed in the quantity and quality of employees is defined as human capital. Therefore, it is important to consider human capital when studying knowledge spillovers. Unquestionably, the flow of skilled labor is one mechanism of knowledge spillover, and the mobility of skilled labor can cause knowledge spillovers.

Many scholars argue that labor mobility is the major way of knowledge spillover, especially tacit knowledge spillover. Zucker et al. (1998) and Almeida and Kogut(1999) showed that when skilled laborers constantly communicate and interact with other people, it can speed up the creation of knowledge and knowledge spillovers, and promote technological innovation. Such phenomenon is more obvious in industrial clusters or cities with large population densities.

In some knowledge-based sectors of the economy, the location and preferences of scientists will affect the geographical position of the innovative activities (Malecki, 1997). Almeida and Kogut (1999) analyzed the mobility of patent holders in the US semiconductor industry, and indicated that the engineer's mobility across firms can affect the local transfer of knowledge, in addition, the knowledge spillovers are embedded in regional labor networks. Fallick et al. (2006) surveyed the computer industry in Silicon Valley and pointed out that the mobility of employees between firms is the main source of regional knowledge spillovers, and the high mobility of employees stimulates the reallocation of talent and resources. In other words, Maliranta et al. (2009) found evidence of the existence of knowledge spillovers across firms, but it is not the most obvious type. In addition, the employees, who had ever engaged in R&D activities in other firms, but now don't engage in R&D activities in the new firms, are able to increase the productivity and profitability of the firms. More specifically, it means that the knowledge embedded in these employees is easy to be copied and imitated without much R&D effort. Furthermore, this paper concluded that the mobility of employees is an important channel of knowledge spillovers.

The empirical research of Filatotchev et al. (2011) on 1,318 high-tech enterprises in Zhongguancun found that returnee entrepreneurs have produced significant knowledge spillover effects and promoted innovation of local high-tech enterprises. The empirical study on Danmark, Stoyanov and Zubanov (2012) provided evidence for the view that employees previously worked in more productive enterprises is related to productivity benefits in the hiring enterprises. Thus, labor mobility is urgent for knowledge spillovers. Song et al. (2003) found that the mobility of employees is able to explain the pattern of patent citation, and indicated that hiring skilled labor from other firms can bring knowledge to the hiring firms. Görg and Strobl (2005) tested whether the owner of domestic firms in Ghana previously worked for MNCs has a correlation with the productivity of firms. The research result shows that firms whose owners worked for MNCs in the same industry immediately prior to setting up their own firm are more productive than other domestic firms. It means that these entrepreneurs accumulate experience working for multinationals and this knowledge can be effectively used in new domestic firms. Similarly, following the trail of labor mobility across firms in Norwegian manufacturing in the period of the 1990s, Balsvik (2011) concluded that firms with high proportions of employees with MNCs experience perform better in productivity. Kim and Marschke (2005) found that with the higher mobility of scientists across firms, firms are more likely to use patents to protect their intellectual property and reduce the loss of firms caused by departing scientists. Kaiser et al. (2015) provided evidence for the view that positive effects of labor mobility stimulate countries' innovation performance.

I.1.3. Measurement of knowledge spillovers

The important role of knowledge spillovers in economic growth is beyond question, but there are two main challenges scholars confront. Firstly, to determine the mechanism or modes of communication that permit knowledge flow. The second challenge is how to identify and measure the knowledge spillovers (Audretsch and Feldman, 2004). Nadiri (1993) conducted an investigation of the results of various countries, periods, and aggregation levels. Overall, the knowledge spillover effect is considered to be very significant and positive, but the estimated effect varied over a fairly wide range. One reason for the differences in effect lies in the methods used to measure spillovers between industries are quite different.

Krugman (1991) emphasized the importance of knowledge spillover and indicated that knowledge spillovers are hard to track. He argued, "Knowledge flows are invisible; they leave no paper trail by which they may be measured and tracked". Therefore, the measurement of knowledge spillovers is the most challenging task.

I.1.3.1. Knowledge production function

The knowledge production function can be used to estimate the impact of technology spillovers on Total Factor Productivity and innovation. Griliches (1979) first proposed the knowledge production function when he analyzed the contribution of R&D to productivity growth and considered knowledge input to be a function of knowledge output. In this model, the observation unit is at the firm level and the firm is assumed to be exogenous and the innovative change these firms are generating is endogenous.

Griliches discussed three main problems of measuring knowledge capital: Firstly, the development process takes time. Due to the lagging structure, current R&D investment may not have a significant effect on output and productivity immediately; Past R&D investment will become partly obsolete and outdated due to depreciation. Therefore, the growth of the net stock of R&D capital investment is not equal to the total level of current or recent R&D investment, that is, there is a problem of depreciation in R&D investment; Lastly, for any industry or sector, its knowledge output does not only come from its R&D investment but is also affected by knowledge spillovers from other industries or sectors. The three questions above point out the direction of research for the estimation of knowledge stock, R&D capital depreciation, and R&D spillover modeling.

Following Griliches's research, economists begin to study how to quantify the extent and impact of knowledge spillovers. Griliches' knowledge production function framework was expanded and empirical research was conducted by Jaffe (1989). He investigated the impact of university R&D activities in 29 states in the United States on the productivity of firms in generating economic knowledge (measured by the number of patents). Jaffe used a modified knowledge production function with a spatial factor that measures the importance of geographic proximity for R&D activities in university and in industry. He used a modified Cobb-Douglas model with two inputs: R&D input by industry, and R&D input by universities. Jaffe's contribution lies in incorporating spatial factors into the knowledge production function, which shifts the research focus on knowledge spillover from the traditional enterprise level to the geographical level.

Based on the knowledge production function developed by Griliches (1979) and Jaffe(1989), many scholars begin to study knowledge spillovers in their framework, and called their model as Griliches-Jaffe knowledge production function. However, the Griliches-Jaffe knowledge production function doesn't distinguish the knowledge spillovers between inter-regions and intra-regions.

In order to explore the most proper unit of observation, scholars begin to consider spatial factors in the knowledge production function. Anselin et al. (1997) further extended this analytical framework using a spatially lagged model, and Fischer and Varga (2003) completely separated the spillover effects of inter-regions and intra-regions in the knowledge production function and considered the time lag in knowledge production. They used university R&D input and firm R&D input as input variables and used patents as output variables to measure the spatial knowledge spillovers in Australia. Greunz (2003) proposed a knowledge production function model, which considers regional geographically and technologically mediated knowledge spillovers together. These findings confirm that the knowledge production function is valid at the geographic level.

We apply the knowledge production function in the case of China in Chapter VI.

I.1.3.2. Paper Trail Approach

The knowledge production function approach is an indirect measure method, while the paper trail approach is a more direct measurement. The paper trail approach uses patent citation to measure knowledge spillovers, which is the mainly used approach in our study. Its main novelty lies in that it can indicate the number of spillovers between sectors, and observe the way in which these spillovers change over time.

In the past two decades, the use of patent citation data has grown dramatically, and patent citations are widely used in the analysis for the reasons as follows: Firstly, citations are an indicator of technological impact and they can be used to prove that the quality of patents increases with the number of citations it receives. Narin et al. (1987) found that the number of citations of a firm’s patents is associated with an increase in firms’ profits. Hence, citation can be used to evaluate the economic value of a firm’s patent. In addition, patent citations can be interpreted as the ‘paper trails’ of knowledge spillover from the inventor of the cited patent to the inventor of the citing patent. Because the patents contain a large amount of data relating to the inventor, the patentee, and their address. Combining the patent citation data with the above-mentioned geographic information can track the knowledge spillover in spaces and institutions over time.

The patent citation also can be divided into backward citation and forward citation according to the citation direction. Forward citations derived from the citations that a patent subsequently receives from other patents, and it is an indicator of subsequent technological impact, and backward citations derived from the citations made by a patent, and it is an indicator of the extent of reliance on previous technology (Harhoff et al., 1999; Henderson et al., 1998). Studies using backward citation information can explore knowledge spillovers between technology classes (Rosenkopf and Nerkar, 2001) or geographic regions (Jaffe et al., 1993; Tijssen, 2001).

Patent citations can reflect the relationship between knowledge and we can use patent citation data to build a knowledge flow network. In addition, the patent document not only contains patent citation information but also the patent applicant, inventor, nationality, technology classification, etc. are provided. Combined with this information, knowledge flow networks between inventors, applicants, within industries or industries, and even between countries and regions can be identified. Empirical research based on patent citations can be carried out at multiple levels of macro and micro perspectives. Patent citation data is extremely rich, spanning a longer period of time and covering almost all industries, meeting the large sample data typically required for overall network analysis.

The knowledge flows and the factors which have an impact on them are not easy to quantify. Compared to measuring the knowledge stock using economic models such as CH and LP, and using Douglas's general production function to measure the total knowledge flows between countries, the method of using patent citation data can explore the mechanism of knowledge spillover. Scholars generally argue that knowledge spillover is an invisible process and its trajectory is difficult to track. Nonetheless, Jaffe et al. (1993) firstly regarded the patent citation as a proxy of the knowledge flow. They developed a matching rate method to test localized knowledge flows by patent citation. They found that patent citations can be used as an “article trail” to measure and track the knowledge spillover. They choose two samples from the U.S. Patent Office’s patent database: one was the 1975 originating cohort containing 950 patents that received 4750 citations by the end of 1989, and the other was the 1980 originating cohort containing 1450 patents that received 5200 citations by 1989. They used patent citations to reflect the direction and intensity of knowledge spillovers. By analyzing these patent citations, the authors discover that knowledge spillovers leave paper trails and these trails are geographically localized; Geographic localization fades over time, and technological areas will affect the localization process, and citations in the same class are likely to be localized. Thompson and Fox-Kean (2005) reassessed the work of Jaffe et al. (1993) using control patents selected under different criteria. In Thompson’s paper, patents use the technology subclass at the six-digit level, which is finer than Jaffe’s technology class at the three-digit level. The results don't support Jaffe's finding of knowledge spillovers are localized at the state and MSA levels, however, it supports Jaffe's finding of significant localization effects at the country level.

Jaffe and Trajtenberg (1999) analyzed the patterns of patent citations taken out by inventors in the UK, France, Germany, the US, and Japan. The results of statistical regression prove that knowledge spillovers are localized. They concluded that patients assigned to the same company are more likely to cite each other; Patents in the same patent class are more likely to cite each other than different classified patents; Inventors who come from the same country are more likely to cite each other than inventors from different countries, And there are clear country-specific citation tendencies, the United Kingdom cites more the US patents than other three countries. Hu and Jaffe (2003) examined the patterns of knowledge spillovers from the U.S.and Japan, Korea and Taiwan, by extracting from the NBER Patent Citations Data File (Hall et al., 2001) all patents taken out in the U.S. by Taiwan, Japan, Korea, and theU.S. from 1963 to 1991. They found that knowledge spillovers from the U.S. and Japan to Korea and Taiwan are in quite different patterns. Korean patents are more likely to cite Japanese patents than US patents, while Taiwan resident inventors learn evenly from both the U.S. and Japanese inventors.

Inspired by the research of Jaffe et al. (1993), Maurseth and Verspagen (2002) used both the patent data and patent citation data from the European Patent Office (EPO) from 1979 to 1996 to study the pattern of knowledge spillovers among countries within the European region. The number of citations between two regions is used as the dependent variable, and treat the geographical distance and technology category as explanatory variables. The results indicate that patent citations are more likely to occur within the same country or between geographically close countries and that geographical distance indeed has a negative impact on knowledge flows and this impact is substantial. The results show that knowledge spillovers are industry-specific and that regions' technological specialization is a significant determinant for their technological interaction. Almeida and Kogut (1997) compared the innovative activity of large and small semiconductor firms, and the analysis of patent citation data shows that small firms are region localized to a greater extent than large firms.

Numerous studies start to apply patent data to measure the direction of knowledge spillovers. Criscuolo et al. (2005) used patent citation data from the EU Patent Office to quantify the relative asset augmenting and exploiting character of foreign-located R&D. Bottazzi and Peri (2003) studied R&D and patent data for eighty-six European Regions, in the period of 1977-1995 and found that spillovers exist for regions within a distance of 300 Km's from each other and spillovers are somewhat weaker, across national borders. Criscuolo and Verspagen (2008) combined the EPO and the USPTO patent citation data to observe knowledge flows. The results show that geographical proximity is not the only variable that impacts knowledge flows, cognitive distance and time are also important. In summary, even though patent citations do not represent knowledge flow in any situation, at the macro level, patent citations can be regarded as an indicator of knowledge spillovers.

Raehel et al. (2011) tested the "home bias" of knowledge flows using patent citations. The results indicate that with the development of international communication, knowledge flows more easily across borders but the situation is different in the more high-tech sectors of ICT and pharmaceuticals, precisely those areas where clusters and agglomeration are believed to be important. Hoekman et al. (2009) studied the research collaboration across 1316 regions in 29 European countries and emphasize the significance of geographical proximity in collaborative knowledge production. Gomes-Cassere et al. (2006) used patent citations in patents granted to the new economy firms in Belgium by the US and the EU Patent Offices to observe the knowledge spillover among firms. The analysis shows that patents between firms also have a localization impact: the closer the distance is, the more frequent the knowledge spillover between firms is, and geographical proximity plays a significant role in promoting the knowledge spillover among firms. Park et al. (2005) improved industrial classification criteria and classifies industries by technical factors. He used patent citation data to investigate distinctively and changing patterns of technological innovation across industries and observe dynamic trends over time. Macgarvie (2005) used patents citations in USPTO taken out by inventors in the United States, Australia, Canada, France, Germany, Italy, Japan, the Netherlands, Switzerland, and the United Kingdom to analyze the determinants of knowledge spillovers between countries, the results indicate that spillovers effects are enhanced by physical and technological proximity and by sharing a common language.

However, this method is questioned by many scholars in terms of validity. Trajtenberg (2001) indicates that patent citation may be a noisy and biased measure of knowledge spillover. Patent citations are regarded as incomplete measures of knowledge spillover for the reason that they can reflect the knowledge spillover process of technologies in the patent documents but cannot represent knowledge contained in other forms. Therefore, we should understand the limitation before we use them.

The main criticism levied against the use of patent citations as ‘flow’ indicators derive from the remark that patent examiners in the US Patent and Trademark Office (USPTO) will add citations to past patents when they think the necessary citations are missing from the inventor's original list of citations (Breschi et al, 2003), and patent examiners, rather than from inventors themselves, are ultimately responsible for the citations attached to patent documents, which is necessary for legal requirements but has no diffusion meaning. The patent examiner may be more likely than inventors to cite related technologies in different industries, and in so doing we confound the effects of the industry with the effects of geography (Thompson, 2006). In addition, in the U.S. many references are not significantly relevant for patentability, and in some situations, we may find reference counts achieve more than 100–200 references (Harhoff et al., 1999; Hall et al., 2000). Besides, there is a strong 'home bias': US patents tend to cite patents granted by the US patent system. Lastly, the USPTO, EPO, and other patent offices also have different patent examination procedures, and this leads to different mechanisms for the citation being added to the patent document.

Although not all inventions are patented, a large amount of patent data ensures that the patent citation data can indicate the overall trend. In addition, since the patent citation itself show the technical value, the cited patent means that it contains a certain technical value, and the existence of the technical value is the premise of the knowledge spillover, which proves the validity of patent citation as a proxy of knowledge spillover. Second, empirical studies prove the validity of patent citations as a proxy for knowledge flow. Jaffe et al. (2000) surveyed a large sample of patent inventors' familiarity with the patents they cited: More than half of the respondents stated that they are familiar with the patent they are citing; Less than one-third of respondents stated that they did not know the patents that they cited, that is, about one-third of patents were added by patent examiners; about 60% of the respondents indicated that they benefited from cited patents through different channels. The survey results show that patent citation data can be interpreted as an effective indicator of knowledge spillover. Alcácer and Gittelman (2006) concluded that examiners are responsible for 63% of citations on the average patent and no evidence shows that the degree of geographic proximity between citing and cited patents differs for inventor and examiner citations. Third, self-citations of patents can be eliminated to some extent. Taking the data of the USPTO as an example, we can use the information about the inventor, institution, and address contained in the patent to exclude reference patents with the same name and the same company or the same name and the same address. In summary, through the reasonable processing of patent citation data, the noise of patent citation data can be reduced, which can show the validity of patent citation data as an indicator of knowledge spillovers.

Chapter II explores the determinants of knowledge spillover in China based on the paper trail method.

I.2. Knowledge spillovers and innovation

I.2.1. Theory of technology, innovation and economic growth

Economic growth has always been an important topic in economics. Acs (2002) pointed out that since knowledge spillovers occur in a spatial context, cities are an appropriate unit of analysis to better understand economic growth.

In the study of neoclassical economists, Marshall (1920) clarified the roles of internal economies and external economies on economic growth. The economy caused by the general development of the industry is called the external economy, while internal economies are dependent on the resources, organization, and operation efficiency of individual enterprises in the industry. Schumpeter (1934) thought economic growth is not caused by external factors, but by the “new combination” of production factors, which is called the “creative destruction” process. It is characterized by the creation of novelty and the destruction of old products and processes.

The Harrod (1939)-Domar (1946) model is the beginning of expressing economic growth theory in terms of models. Since then, the economic growth theory applicate mathematical tools in research. The Harrod-Domar model highlights the role of “capital accumulation” in economic growth, and there are two main factors determining the level of economic growth of a country: the savings ratio and the capital-output ratio. Since the capital-output ratio remains constant in the short run, thus, the growth rate of a country depends mostly on the savings ratio. The most important assumption in the model is that the production is under the condition of fixed proportions, and in such a situation, it is not possible for labor to substitute for capital in production. The limitation of the model is the ‘knife-edge’ properties. The steady-state growth was unstable and it is in the "unstable equilibrium", and the reason is that the parameters that determine the equilibrium condition are all exogenous. Because of a certain element, changing one of the parameter values without changing other parameter values, the economy will deviate from equilibrium and can’t return to equilibrium.

Based on the criticism of the Harrod-Domar model, Solow (1956) proposed a neoclassical growth model, and he thought that the steady growth of full employment could be achieved by adjusting the ratio of labor to capital in production through the capitalist market mechanism. Swan (1956) proposed a similar model. They revised the Harrod-Domar model to overcome the defect of the "knife-edge" property, and they accepted all the Harrod-Domar assumptions except that of fixed proportions. Solow-Swan model satisfies the general equilibrium condition and becomes representative of the neoclassical growth theory. In the Solow-Swan model, three new hypotheses were introduced: (1) the production function used in the model is in a neoclassical form, therefore, capital and labor can be surrogated for each other; (2) the labor supply and capital stock are in balance; and (3) fixed depreciation rate.

In the Harrod-Domar model, technological improvement is represented by the improvement of labor productivity. In the Solow-Swan model, the technological improvement is represented by the change of an independent variable in the aggregate production function, which is a well-known variable A, representing technology level.

Since the main feature of the neoclassical production function is the diminishing marginal return of input factors, the long-term per capita economic growth rate tends to zero in the absence of technological improvement. Therefore, in the neoclassical economic growth model, the long-run sustained economic growth can only rely on exogenous technological improvement, which is assumed as a public good.

Following closely with Solow's (1956) growth model, Solow (1957), in his article "Technical Change and the Aggregate Production Function", proposed a method of aggregate production function analysis and applied this method to validate the neoclassical model. He found that capital and labor can only explain about 12.5% of total output. Therefore, Solow used exogenous "Residual" to explain the technological change, thus explaining 87.5% of total output. He believed that technological change is a more important economic growth factor than labor and capital.

The limitation of the neoclassical economic growth model lies in that economic growth rates depend on variables outside the model and cannot be explained in the model. The neoclassical school believed that when the market fails in terms of supply and demand for technological innovation, or when technological innovation resources cannot meet the requirements of economic development. The government should immediately adopt indirect regulation and control measures such as finance, taxation, law, and government procurement to intervene in technological innovation activities to further enhance the role of technological improvement in economic development.

The Solow-Swan model has always dominated the development of economic growth theory. However, due to some shortcomings of the model, in the mid-1980s, this model was gradually replaced by the endogenous economic growth model. But the Solow-Swan model is the starting point for almost all research on economic growth issues.

Considering the limitations of the neoclassical economic growth theory, Arrow (1962) published the article "The Economic Implications of Learning by doing" in The Review of Economic Studies. It is a meaningful paper, which is regarded by the economics industry as the forerunner of the theory of technological endogenous economic growth. Arrow believed that although Solow has made a lot of analysis on economic growth theory and regards technological change as the main driving force for economic growth, it is doubtful that Solow treats technological progress as an exogenous variable. Because exogenous variables do not explain the dynamics and processes of economic growth.

Arrow proposed that technological change or productivity improvement is a by-product of capital accumulation. Knowledge is generated within the production process. He defined the way in which knowledge is accumulated through investment activities as learning by doing. He believed that not only manufacturers can increase productivity by accumulating production, but other manufacturers can also increase productivity by “learning”, that is, non-competitive knowledge has externalities. Moreover, the knowledge accumulated in -/+-*- learning by doing will grow indefinitely with the expansion of the investment scale, which provides a guarantee for long-term economic development. The technological change that is regarded as a godsend in the neoclassical model is not something else, but the knowledge accumulated through learning by doing. In this sense, technological change can be regarded as an endogenous variable determined by the economic system. However, in Arrow's “learning by doing” model, the technological change rate of a society ultimately depends on the exogenous population growth rate. Then the model doesn’t solve the problem of “Solow residual”, that is to say, how to make the technological change endogenous.

I.2.2. Endogenous growth theory

In the 1980s, with the groundbreaking contributions of Romer (1986) and Lucas (1988), which draw in part on work by Arrow (1962), Uzawa (1965) and Sheshinski (1967), a large body of literature on endogenous economic growth has arisen. These literature introduce factors such as increasing returns to scale, imperfect competition, and human capital into the growth model. They give explanations of the economic performance in different countries at different stages of development. The endogenous economic growth theory introduces factors such as knowledge and human capital into the economic growth model. The endogenous growth theory breaks through the assumptions in the neoclassical economic growth theory that the factor returns are diminished or the factor returns are constant and explain the source and motivation of economic growth. Technology is no longer an exogenous, uncontrollable thing, but a product of human investment for its own benefit. The endogenous growth theory affirms the important role of government intervention in economic growth and believes that the equilibrium growth rate of economic competition is usually lower than the social optimal growth rate of the economy. In the endogenous growth theory, technological knowledge becomes a partly private and partly public good (Romer,1990; Grossman and Helpman, 1991).

Conceptually, Roberts and Setterfield (2007) gave three definitions of endogenous growth theory: (1) The growth rate is determined by the growth model itself and is determined by the endogenous variables of the model rather than resorting to exogenous variables. (2) Incorporate technological change into the growth model rather than regarding it as a gift from God. (3) From the evolutionary point of view, it only takes time as the only exogenous variable, and regards growth as the path dependence of the previous growth history.

Following Arrow's method of using technological externality or knowledge spillovers to explain economic growth, Romer (1986) built a model which regards technological change as endogenous variable, and the model is based on an understanding of the externalities of knowledge. He assumed that knowledge is a factor of production, and that the manufacturer can gain knowledge in the process of capital accumulation. Knowledge is different from common commodities in that knowledge has a spillover effect, which enables the knowledge produced by any manufacturer to increase the productivity of the whole society. Due to the existence of knowledge spillovers, the marginal productivity of capital will not decrease indefinitely, and the private yield of firms will be lower than the rate of social returns.

Spatial knowledge spillovers based on regional knowledge stocks and knowledge absorption capacity and human capital stocks have become the main source of economic growth. Romer suggested that the government could provide subsidies to manufacturers of production knowledge, or tax other production while subsidizing knowledge production. These policies can motivate private firms to produce knowledge that will increase economic growth rates and social welfare levels. He put the production function of a manufacturer as：

$$Y = F(K, K_{\mathrm{a}}, L)$$

Comparing with the neoclassical production function, it introduces one more variable K_{a}. K and L is the manufacturer's own capital and labor input, K_{a} is the average input of other manufacturers. As for the manufacturer, it can't determine the amount of input from others, but the amount of input from others will affect its productivity because of its positive externalities.

Lucas(1988) proposed the human capital accumulation theory and introduced human capital as an independent factor into the Solow model, considering that the external effects of human capital play a central role in economic growth, extending from individuals to others, and contributing to the productivity of all factors of production. He described two types of human capital models: a model emphasizing capital accumulation through schooling, and a model emphasizing specialized human capital accumulation through learning-by-doing. The knowledge embedded in human capital is assumed to be an input in production, which increases marginal productivity through school education and learning by doing. Lucas believed that the human capital accumulation happens when the labor is in leisure time, while leisure time is limited. When this person dies, the human capital disappears at the same time, and the human capital accumulation activities also stop. But the knowledge that this individual possesses, including all non-competitive products he produces, such as scientific theory, technical principles and patents remain after his death. Lucas put the production function is as:

$$y = Ak^{a}(uh)^{1-a} h_{a}^{\beta},$$

A is the technical coefficient, k is the physical capital invested by each worker, h is the input of human capital, and u is the time a worker uses for production. Suppose a worker with skill h devotes the fraction $u(h)$ of his non-leisure time to current production, and the remaining 1-$u(h)$ to human capital accumulation, and h_a is the human capital with external effect.

Chapter II: Knowledge Flows Within Chinese Administrative Provinces: A Patent Citation Analysis

II.1. Introduction

The objective of this chapter is to analyze the determinants of knowledge flows between Chinese administrative Provinces. This work is one of the first attempts to apply economic analysis of R&D knowledge spillovers in developing countries, and it explores empirically the determinants of patent citation in China at regional level. Most empirical research about knowledge spillovers concentrate on developed countries and the number of studies on China, although growing, is still modest. In addition, most studies that look at knowledge flows in China focus on the firm level (Xiang et al., 2013; Hansen and Hansen, 2020) or try to estimate the impact of knowledge flows from abroad (Qiu et al., 2017). To our knowledge, very little study so far has attempted to understand the determinants of inter-regional knowledge spillovers in China. Those who did found that intraregional and international collaborations are the main channels of knowledge exchange in China, while inter-regional knowledge exchange is relatively weak (Gao et al., 2011). It might therefore be important to understand the determinants of inter-Provinces knowledge flows in China in order to develop policies that could unlock the existing impediments.

Focusing on knowledge flows between Chinese Provinces enables to examine standard research questions such as the effect of the geographical distance or technological proximity between two Provinces. But, and more originally, it also enables to explore to which extent the characteristics of the Provinces affect knowledge flows. In particular, in this chapter, we focus on the effect of the research structure of each Chinese Province. We use factor analysis in order to measure the private versus public research intensity of each Chinese administrative province, thus introducing a distinction between Provinces that are more oriented towards public versus private research. This is in line with research that go beyond the geographic or technological effect and look at the institutional dimension in order to explain knowledge flows (Gittelman, 2006; Rodríguez-Pose, 2013).

We rely on patent citations to measure and capture knowledge spillovers (Jaffe and de Rassenfosse, 2017). For several decades, economic studies have used patent statistics to measure innovative activity[1]. Since the 1990s, patent citations have also been widely used to measure knowledge spillovers (Jaffe et al., 1993)[2]. Patent citations, however, remain an imperfect measure of knowledge externalities. First, because they only reflect knowledge codified in the patent document and neglect knowledge contained in other forms. Second, because patent examiners often add citations during the examination process, thus inducing bias and noise in the measure of knowledge flows (Jaffe et al., 2000; Alcacer and Gittelman, 2006; Alcacer et al., 2009; Lampe, 2012; Roach and Cohen, 2013; Moser et al., 2018; Corsino et al., 2019). Despite this important caveat, for Jaffe and de Rassenfosse (2019) there is today a broad consensus on the fact that patent citations remain an essential tool to measure inventions' characteristics (their impact, value, generality, originality, etc.) and knowledge flows between economic actors.

In line with this literature, we use data on the patent inventors' location to identify the geographic location of the cited and citing patents. Each citation between two patents is therefore associated to two of the 31 Chinese administrative Provinces (the citing and the cited Province). Relying on patent data from the US Patent and Trademark Office during the period 1995-2019, we are able to identify 27 118 patent citation pairs.

Our econometric results show that, as expected, geographical and technological distance between Provinces are negatively correlated with patent citations. In addition, we find that Chinese administrative provinces that show bigger intensity with regard to public and to

---

[1] As acknowledged by Griliches: "In spite of all the difficulties, patents statistics remain a unique resource for the analysis of the process of technical change. Nothing else even comes close in the quantity of available data, accessibility, and the potential industrial, organizational, and technological detail" (Griliches, 1990, p. 1702)

[2] The use of patent citation is expected to offer a solution to the challenge put forward by Paul Krugman: "knowledge flows are invisible ; they leave no paper trail by which they may be measured and tracked and there is nothing to prevent the theorist from assuming anything about them that she likes" (Krugman, 1991, p. 53-54).

private research both receive more patent citations and cite more other patents in other Provinces. This effect is particularly significant when the two regions are specialized in private research, thus suggesting that two Chinese Province strongly oriented towards private research are more likely to generate knowledge inflows and outflows one with each other. On the other hand, this effect is not significant when the two Provinces are specialized in public research. Finally, and at odds with prior research (Zucker et al., 1998; Breschi and Lissoni, 2003; Morrison, 2008; Frenken et al., 2010; Ott and Rondé, 2019), we find that controlling for social proximity between Provinces does not impact the effect of geographical and technological proximity. These results, if confirmed by further studies, might have significant policy implications as to the influence of Chinese Provinces research structure on knowledge circulation within China.

The remainder of the chapter is structured as follows. Section 2 surveys the literature related to the determinants of knowledge spillovers. Section 3 details the empirical design of the research and provides information as to the data we use. Section 4 reports our main empirical findings. Section 5 concludes.

II.2. Related literature on the determinants of knowledge flows

Since the seminal work of Griliches (1957) and Jaffe (1986), the economic literature has highlighted a number of factors determining the intensity of knowledge spillovers. These factors can be grouped into four categories: geographic and national proximity (linked to cultural and linguistic proximity), technological proximity, temporal proximity, and social proximity (social network).

The geographic dimension is the one that has received the most attention. Extending the work of Jaffe (1989), the paper of Jaffe, Henderson, and Trajtenberg (1993) is generally considered as pioneer in the analysis of the geographic dimension of externalities using patent citations. Starting from a set of patents filed by American universities and firms and citations made to these patents, and comparing this sample of citing patents to a control group, they show that geographic proximity does indeed play a significant role in the probability to cite a patent. This is true at the national level but also at the regional level (citations have a higher probability of being made by a patent filed in the same country, in the same state, and even in the same metropolitan area).

If the methodology used by Jaffe et al. (1993) was criticized by Thompson and Fox-Kean (2005), who obtained less clear-cut results (although without completely questioning the importance of the geographical dimension), the vast majority of studies that have followed highlighted a significant geographic effect of knowledge spillovers. For example, Maurseth and Verspagen (2002), on European data, find that “Geographical distance has a negative and substantial impact on knowledge flows”. Likewise, Belenzon and Schankerman (2013) find that “Citations to patents decline sharply with distance and are strongly constrained by state borders”. Finally, Murata et al. (2014), revisits “the debate between Thompson and Fox-Kean and Henderson, Jaffe, and Trajtenberg on the existence of localized knowledge spillovers and find solid evidence supporting localization”. Interestingly, we also have evidence that progress in ICTs does not seem to lower the importance of geographic proximity for knowledge flows, on the contrary (Sonn and Storper, 2008).

Geographical distance reduces the flow of knowledge because it limits the circulation of tacit knowledge. Geographic proximity enhances the possibility of economic agents exchanging ideas and, particularly, tacit knowledge. Codified knowledge might easily circulate despite distance but tacit knowledge requires human interactions that, despite progress of ICTs, remain largely localized. This is why, as famously stated by Feldman (1994): “knowledge crosses more easily corridors and streets than oceans and continents”.

In relation to geographic proximity, most studies also highlight a country effect for knowledge spillovers. For example, Jaffe and Trajtenberg (1999) analyze patent citations between US, UK, France, Germany, and Japan and find that: “patents whose inventors reside in the same country are typically 30 to 80% more likely to cite each other than inventors from other countries, and these citations come sooner”. This result is confirmed by Branstetter (2001) who shows with data on the US and Japanese firms that knowledge spillovers are primarily intra-national and by Maurseth and Verspagen (2002) who find that the country effect remains even if regions share the same language. In the case of China, Qiu et al. (2017) find that only companies in the most advanced Chinese regions benefit from international externalities and that domestic collaboration has shown a larger positive impact on corporate innovation than international collaboration in recent years. These results suggest an effect of national institutions that is different from the simple effect of geographic proximity.

Following Jaffe (1986), the literature has also highlighted the importance of technological proximity to explain the flow of knowledge between economic actors. For example, Autant-Bernard (2001) shows, using French data, that technological proximity seems to play a more important role than geographical proximity. She also highlights the importance of human mobility in the circulation of knowledge, a factor to which we will come back. Maurseth and Verspagen (2002, p. 531) also show that while patent citations in Europe are affected by geographic proximity: “patent citations are industry specific and occur most often between regions that are specialized in industrial sectors with specific technological linkages between them”. Orlando (2004) obtains similar results on company data.

The importance of technological proximity is primarily explained by the fact that companies working on the same research projects have more to learn than companies working on very different things. The potential of knowledge transfer is thus much greater when the companies are technologically close. This had already been underlined by Griliches (1991): “The photographic equipment industry and the scientific instruments industry may not buy much from each other but may be, in a sense, working on similar things and hence benefiting much from each other’s research”. In addition, companies working in technologically related fields also find it easier to absorb knowledge (Cohen and Levinthal, 1989). It is indeed easier to understand and reuse knowledge developed by its technological neighbors.

The time dimension is also an important determinant of knowledge flows. Caballero and Jaffe (1993) and Jaffe and Trajtenberg (1999) have developed a model for the diffusion of knowledge over time based on two opposite effects: first, the more time passes, the more knowledge is likely to diffuse, which thus mechanically increases its potential for future dissemination. Second, the more time passes, the more knowledge becomes obsolete, which reduces its potential for future dissemination. These two opposing forces thus make it possible to envisage an optimum of diffusion over time, that is to say, the moment when a piece of knowledge has reached its peak of maximum diffusion. This diffusion peak obviously varies according to the country, sector, or type of organization. Jaffe and Trajtenberg show, for example, that the diffusion peak is reached more quickly in electronics than in chemistry or mechanics (i.e., technologies are obsolete more quickly in electronics). More recently, Mehta et al. (2010) using a modified method find that the diffusion peak arrives earlier than expected by Jaffe and Trajtenberg, but do not significantly question their result. Finally, Bacchiocchi and Montobbio (2009) use the same methodology to study the impact of knowledge-producing organizations on the dissemination of knowledge. They show that the technologies developed by universities diffuse more quickly than those developed by firms.

Finally, several studies have affirmed the importance of social network, that is to say of the previous existence of links between economic actors, to explain knowledge flows. This literature directly questions Marshall's famous quote that "knowledge flows in the air". For Breschi and Lissoni (2003, po. 6) "this view provides a naive portrait of the channels along which knowledge flow". Of course, knowledge does not circulate in the air but between individuals. Consequently, social proximity between individuals is an essential determinant in explaining the flow of knowledge. For instance, Almeida and Kogut (1999) show that engineers' inter-firm mobility influences the local transfer of knowledge. Also, Agrawal et al. (2006) use data on inventor mobility to highlight the importance of social proximity to explain knowledge flows. They find that the flow of knowledge between an inventor's departure place and arrival place is 50% higher than if the inventor had not lived in the departure place. The overall evidence is therefore consistent with a view that social relationships, not just physical proximity, are important in determining observed patterns of knowledge diffusion.

Some studies even suggest that the effect of social proximity is so strong that the geographic effect disappears when properly controlled for social proximity. For instance, Breschi and Lissoni (2003) find that: "localization effects tend to vanish where citing and cited patents are not linked to each other by any network relationship. On the contrary, knowledge flows, as evidenced by patent citations, are strongly localized to the extent that labor mobility and network ties also are. We interpret these results as evidence that geography is not a sufficient condition for accessing a local pool of knowledge, but it requires active participation in a network of knowledge exchanges". Rondé and Hussler (2005) are in the same line when demonstrating that "the impact on regional innovation of unintended knowledge flows decrease when voluntary actions (deliberate interactive competences building) are introduced in the model" and Singh (2005) goes also in the same direction: "The existence of a network tie is found to be associated with a greater probability of knowledge flow. Furthermore, the effect of regional or firm boundaries on knowledge flow decreases once interpersonal ties have been accounted for. In fact, being in the same region or firm is found to have little additional effect on the probability of knowledge flow among inventors who already have close network ties".

In summary, a very large number of studies have explored the determinants of knowledge flows between economic actors. These flows are influenced by geographic proximity (possibly through social proximity which is constrained by geographic proximity), technological and temporal proximity. Also, the observation of a significant national effect seems to indicate that institutional proximity plays a role (Boschma, 2004). In the rest of this article, we focus on the determinants of knowledge flows between Chinese Provinces. Besides the geographical and technological effect, we are interested in the effect played by the research structure of the Provinces. As a matter of fact, when considering knowledge flows between economic actors, the magnitude of these flows depends not only on different forms of proximity between actors, but also on their capacities to emit and absorb knowledge. Since the work of Cohen and Levinthal (1989), we know that these capacities depend on the research investment of the different actors. Moreover, from a more macroeconomic perspective, investments in local institutional research and educational structures in order to higher the quantity and the quality of human capital furthermore determine the absorption of knowledge and innovation generated elsewhere (Barro, 1996; Rodriguez-Pose and Crescenzi, 2008). Applied to our Chinese regions, this means that we must consider the research structures of the different Provinces. Indeed, Rondé and Hussler (2005) have shown the positive effects of the scientific density of territories (Nuts 3 level) on the innovation capacities of firms via the development of relational capacities in order to generate and absorb knowledge flows. In an interesting study focusing on the biotechnology industry, Guittelman (2006) shows that research institutions have differentiated effect on knowledge flows according to their organizational structure which suggests that specific combinations of individuals' knowledge with organizational capabilities matter to explain innovation outcome. "Technological performance is explained by the heterogeneity of organizations and individual relationships within countries, rather than an overall institutional effect that operates similarly upon all organizations in the countries" (Guittelman, 2006, p1067). In order to take into account the role of regional organizational research structures, we contrast the Provinces specializing in public research with those specializing in private research and analyze the consequences of this specialization on the flow of knowledge. To our knowledge, this effect has not been measured in the literature yet. However, it is likely that it reflects a form of institutional and/or cognitive proximity (Boshma, 2005; Boschma and Frenken, 2010) between the actors and therefore ultimately has a significant impact.

II.3. Empirical design

II.3.1. Data collection

We conduct our analysis of knowledge flows at the level of Chinese Provinces. We have therefore collected detailed information on 31 out of the 34 Chinese administrative Provinces. Taiwan, Hong Kong and Macao have been excluded from our analysis due both to their very specific situation and to missing information. The period considered is 1995-2019.

The data used in this study comes from two main sources: First, we use data provided by the China Statistical Yearbook, prepared by the National Bureau of Statistics of China. This yearbook displays every year detailed information as to the economic activity and performance of the Chinese Provinces. Second, information about patents and patent citations has been provided by the US Patent and Trademark Office (USPTO). We used PatentsView[3] to collect information about patent documents and, in particular, patent citation data. In detail, we downloaded the patent and patent citation data on the website of PatentsView by searching with the field "Inventor Country equals CN", and the information on patents granted during January 1995–December 2019 was able to be retrieved. The USPTO lists the names of all inventors on patents. The addresses (city and county) are listed for all inventors on patents. We use the inventor city to assign individual patents to their corresponding Provinces. Following the first-named inventor assignment principle (Acs et al., 2002; Bottazzi and Peri, 2003; Hu and Jaffe, 2003; Singh, 2008; Stolpe, 2002; Trajtenberg, 2001), we manually distinguish which Province the patents belong to by the address of the first inventor. After the process of cleaning the data, we further combine the citation pairs at the provincial unit level. We use USPTO citations to build a set of matrices that map citations between any two Provinces. Each cell of the matrix is the number of citations in patents with the first inventor resident in a Province to patents with the first inventor resident in another Province.

We focus on all the utility patents granted by the USPTO to the first inventors residing in China over the 1995–2019 period. Furthermore, we use patent citations to identify all the citations made between patents with the first inventors residing in China. In the considered period, we have a total of 27118 patent citation pairs that we can then assign to Chinese Provinces, based on the first inventors' locations of citing and cited Provinces.

We focus on Chinese patents registered in the US patent office for three reasons. First, there is no patent citation information available in any of the Chinese patent databases, which makes it impossible for us to get patent citation data of the patents granted by the China

[3] Patents View is a patent data visualization and analysis platform intended to increase the value, utility, and transparency of US patent data. The initiative is supported by the Office of the Chief Economist in the USPTO, with additional support from the U.S. Department of Agriculture (USDA). (www.patentsview.org)

National Intellectual Property Administration (CNIPA) and undertake an in-depth patent citation analysis. At the same time, the US patent database offers detailed information on patent citations, and knowledge spillover can be analyzed by patent citation data. Second, since the USPTO is the world's highest-level patent organization and since the US is China's most important destination country for patent applications overseas, focusing on USPTO data allows us to keep enough patents in our database. Third, there are certain studies that have used the USPTO database to assess knowledge spillovers in China (Yu and Wu, 2014; Xiang et al., 2013; Gao et al., 2011; Wu and Mathews, 2012), which supports our choice of the USPTO to handle our case in China.

Considering newer patents are referenced less frequently and there were fewer USPTO applications in the early years, we decided to take a cross-section of the entire period rather than breaking it up into separate periods (Maurseth and Verspagen, 2002). As a result, we are unable to evaluate the trends in data.

II.3.2. The public versus private research structure of Chinese Provinces

Since we seek to understand the effect of Provinces' research structures on knowledge flows, we use the factor analysis method in order to characterize the research structure of each Chinese Provinces. Factor analysis is a multivariate statistical analysis method that converts multiple variables into a small number of factors that summarize the explanatory capacity of the original variables. The main idea is to reduce the dimensions and simplify the data in order to focus on a small number of common factors (Buesa et al, 2010; Jellema and Roland, 2011). In order to take into account the quality and quantity of institutional research structures, we chose an initial number of 20 variables extracted from the Statistical Yearbook of China in 2016. The factor analysis led to three significant common factors that are indicated in Table 1:

an indicator of public research intensity (F1), an indicator of private research intensity (F2), and an indicator of transportation infrastructure quality (F3). The cumulative variance contribution rate reaches 90.64%, thus exceeding the general requirement of 85%.

Each province can then be characterized according to its performance related to F1, F2, and F3. Figure 1 indicates how the 31 Provinces perform along the F1 and F2 dimensions. We see that some Provinces are quite intensive in public research but much less in private research (for instance Beijing). Some Provinces, conversely, are quite intensive in private research but much less in public research (for instance, Zhejiang). Some are good at both public and private research (for instance, Guangdong). The bulk of the Provinces remain below zero both for public and private research.

Using each Province scores on F1, F2, F3, and the co-variance contribution rate of each common factor, we can calculate the comprehensive innovation scores of 31 Provinces, as shown in Appendix I. We can see that Beijing, Jiangsu, Guangdong, Sichuan, Shanghai, Shandong, and Zhejiang are province-level administrative regions with the strongest regional innovation capability. Most of these Provinces are located in the eastern region, which has a large number of famous firms and universities and a better innovation environment than the central region and western region. We see also that a majority of Provinces have a negative score, indicating innovation in China is concentrated in a minority of Provinces. Using the factor analysis method, we have therefore a preliminary understanding of China's regional innovation capability. In the next part, we will introduce the scores of F1 and F2 of each province-level administrative region as the independent variables into the estimation model of knowledge flows[4]

[4] The factor analysis has been conducted for different years, at the beginning, the middle and the end of our sample. The time does not affect the outcome of the factor analysis. For this reason, in our econometric treatment we retain the factor analysis performed at year 2016 and displayed in this section.

Table 1 Factor analysis.

Component	F1	F2	F3
Papers published by the scientific research institutions	0.978		
R&D personnel full-time equivalent in the scientific research institution	0.962		
Publishing scientific works in the scientific research institutions	0.961		
R&D expenditures in the scientific research institutions with regard to GDP(%)	0.940		
Number of R&D projects in scientific research institutions	0.948		
Number of Scientific research institutions	0.853		
Scientific Books Published in the University	0.780		
University R&D expenture with regard to GDP (%)	0.786		
Number of R&D projects in the University	0.670		
Published scientific papers of the University	0.667		
R&D personnel full-time equivalent in the university	0.649		
R&D staff full-time equivalent in industrial enterprises above designated size		0.983	
Sales revenue of new products of industrial enterprises above designated size		0.975	
Number of new product projects of industrial enterprises above designated size		0.970	
Number of R&D projects of industrial enterprises above designated size		0.968	
Funds for developing new products of industrial enterprises above designated size		0.956	
Firms R&D expenture with regard to GDP(%)		0.765	
Railway business density			0.854
Long-distance optical cable line density			0.806
Highway density			0.772

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization

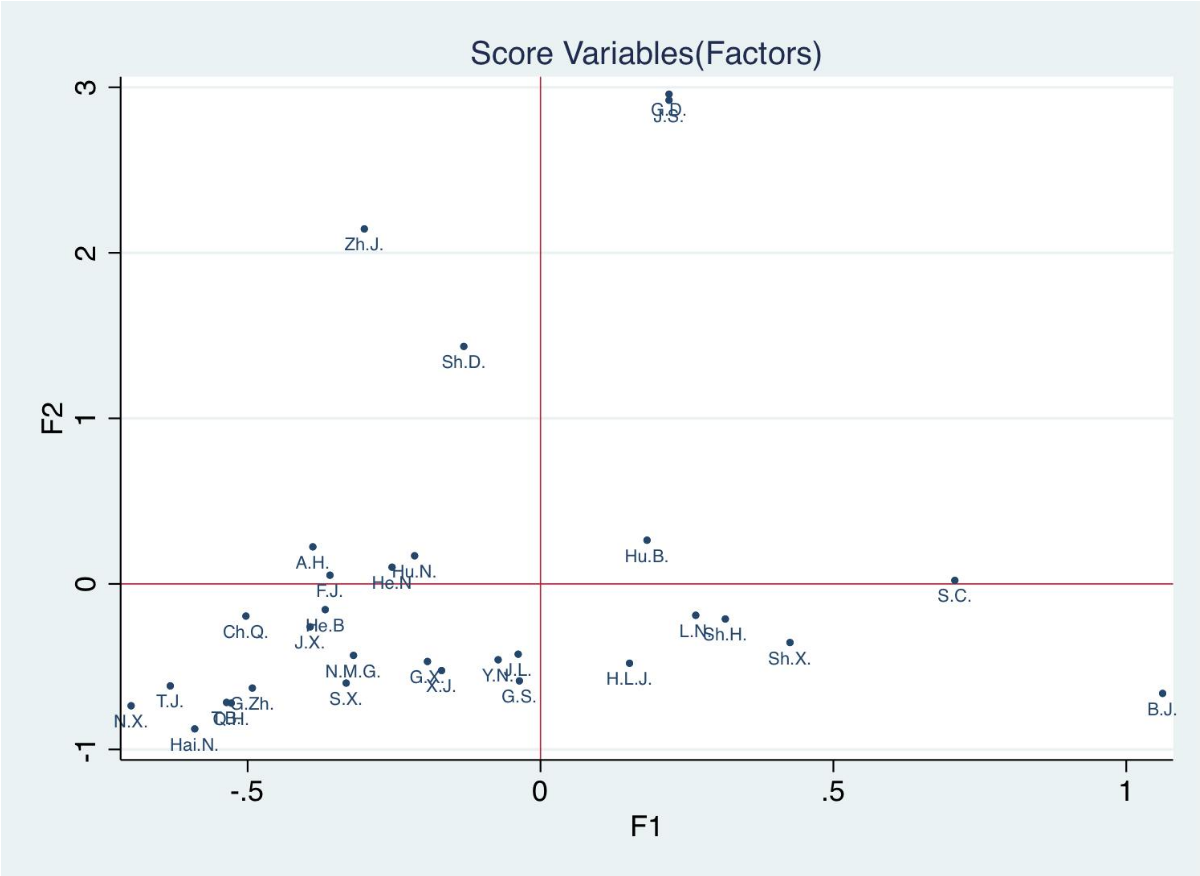

Figure 1. Factor 1 and Factor 2 scores of 31 provinces (The meaning of the abbreviation is in Appendix I).

II.3.3. Econometric treatment

Since we are interested in the determinants of knowledge flows between Chinese Provinces we rely, as dependent variable, on the number of citations C_{ij} between Province i and Province j. Please note that we pay attention to the direction of the citations, which means that C_{ij}, which is the number of citations that Province i made to Province j's patents, is different from C_{ji}, which is the number of citations that Province j made to Province i's patents. In other words in C_{ij}, i is the citing Province and j the cited one. Therefore, as we consider 31 Provinces in our database, our econometric model is based on 31 * 30 = 930 observations.

In order to explain the number of citations between two Provinces, we rely on a set of independent variables that are summarized in Table 2 and Table 3. First, and in line with the literature on the determinants of knowledge spillovers, we focus on geographical and technological distance. The geographical distance, denoted by D_{ij}, is great circle distance (in kilometers) between the regions of the citing and cited inventors, based on coordinates for the regions or their geographical centers.

Technological proximity, noted S_{ij}, is calculated by using the international patent classifications of patents in region i and region j. Following Jaffe (1986), we use the distribution of the regions' patents over 8 patent classes to calculate the technological proximity between two regions. Jaffe's formula was originally used to measure the proximity of the technological structure between firms. But it can also be used to capture the extent of technological overlap between province-level administrative units pairs.

Second, we also introduce as an independent variable the public versus private research intensity of each Province. This is done by introducing the F1 and F2 scores calculated in section 3.2. F1 indicates the score of public-sector R&D, and F2 is the score of the private-sector R&D. These scores capture differences into regional research structure. Some regions are more intensive in public research and others more intensive in private research. Moreover, as figure 1 delimitates four zones (F1 > 0 and F2 > 0, F1 > 0 and F2 <0, F1< 0 and F2 > 0, F1< 0 and F2 > 0), two of them are particularly remarkable, namely, the regions intensive in private research and the one intensive in public research. Therefore, it could be interesting to know whether the transmitters and receivers of knowledge comes from the same categories of regions (public vs private R&D intensive) or not. To check for this point, we introduce three dummies variables D1, D2 and D3, as detailed in Table 2.

As control variables, we include geography dummies for each cited and citing region. We also use the World Bank's standards to measure the degree of economic development. According to this criteria, the 31 provincial administrative regions are divided into six economic regions: Northeast, Bohai, Southeast, Central, West, and North[5]. Finally, we use co-inventors measure in order to capture the social proximity (social network effect) between regions. In doing so, we want to contribute to the debate "unintended spillovers vs deliberate interactions" debate which, to our knowledge, has only been informed by empirical studies from American or European data. The number of co-inventors is this defined as the frequency of collaboration between inventors in region i and inventors in region j.

We express the independent variables of geographical distance, number of patents as natural logs in order to lessen the impact of outliers, and to reduce heteroscedasticity. Since our dependent variable is a discrete and positive variable with overdispersion, we use count

[5] The Northeast includes Heilongjiang, Jilin and Liaoning; The Bohai includes Beijing, Tianjin, Hebei and Shandong; The Southeast includes Shanghai, Jiangsu, Zhejiang, Fujian and Guangdong; The Central includes Henan, Hubei, Hunan, Anhui and Jiangxi; The Southwest includes Chongqing, Sichuan, Yunnan, Hainan, Guizhou and Guangxi; The Northwest includes Shanxi, Shaanxi, Gansu, Ningxia, Inner Mongolia, Xinjiang, Qinghai and Tibet. There are 5 fixed-effect dummies for citing regions and for cited regions, respectively.

data models such that Negative binomial regression and Tobit regression (Hausman, et al. 1984).

Table 2 Description of variables included in our model.

Variable	Definition
Dependent Variable	
C_{ij}	The number of patent citations between provincial-level administrative unit *i* and *j*.
Independent Variable	
$Co\text{-}patenting_{ij}$	The number of patent cooperation between provincial-level administrative unit *i* and *j*.
P_i	The total number of patents in provincial-level administrative unit *i*.
P_j	The total number of patents in provincial-level administrative unit *j*.
D_{ij}	Geography distance, the great circle distance (in kilometers) between the provincial-level administrative units of the citing and cited inventors, based on coordinates for the regions or their geographical centers.
S_{ij}	Technological proximity between provincial-level administrative unit *i* and *j*.
$Traffic_i$	The score of the traffic of region *i*.
$Traffic_j$	The score of the traffic of region *j*.
$F1_i$	The score of public-sector R&D performance of region *i*.
$F2_i$	The score of the private-sector R&D performance of region *i*.
$F1_j$	The score of public-sector R&D performance of region *j*.
$F2_j$	The score of the private-sector R&D performance of region *j*.
$F1_i \times F1_j$	Interaction of the score of public-sector R&D performance between region *i* and *j*.
$F2_i \times F2_j$	Interaction of the score of private-sector R&D performance between region *i* and *j*.
D1	Dummy variable taking value 1 if the score of the public-sector R&D performance of region *i* and *j* both are above 0.
D2	Dummy variable taking value 1 if the score of the private-sector R&D performance of region *i* and *j* both are above 0.
D3	Dummy variable taking value 1 if the score of the public-sector R&D performance of region *i* (*j*) is above 0, and the score of the private-sector R&D performance of region *j* (*i*) is above 0.
Geography dummy variables	Geography dummy variables for each cited and citing region, include 6 regions of Northeast, Bohai, Southeast, Central, West, and North. 5 dummy variables for citing regions, and 5 dummy variable for cited regions.

Table 3 Descriptive statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
C_{ij}	930	7.113	44.64	0	762
$Co\text{-}patenting_{ij}$	930	18.73	99.36	0	1,328
S_{ij}	930	0.780	0.179	0	0.997
lnP_{ij}	930	11.23	3.211	1.099	20.41
lndis	930	14.02	0.597	11.54	15.17
$F1_i$	930	0	0.984	-0.699	5.062
$F2_i$	930	0	0.984	-0.876	2.958
$F1_j$	930	0	0.984	-0.699	5.062
$F2_j$	930	0	0.984	-0.876	2.958
$Traffic_i$	930	0	0.984	-1.345	3.550
$Traffic_j$	930	0	0.984	-1.345	3.550
D1	930	0.0774	0.267	0	1
D2	930	0.0968	0.296	0	1
D3	930	0.172	0.378	0	1

II.4. Empirical results and discussion

All the estimation results of regression models are shown in Table 4 and Table 5. In Table 4, the Tobit model uses the logarithm of the citation count as the dependent variable, Model 1 and Model 2 reports the result of Tobit regression. Model 3 and Model 4 show results of the negative binomial estimates. Results of the generalized negative binomial model are reported in Model 5 and Model 6. Table 4 reports the results with the F1 and F2 variables and with interaction terms whereas Table 5 reports results with the dummy variables D1, D2 and D3. The overall results are generally consistent with our hypotheses, suggesting that geographical distance, technological proximity and the research structure of Provinces have significant impact on knowledge spillovers in China.

From Table 4, we observe that the coefficient of the geographical distance variable is negatively and highly significant in all the specifications, and the absolute value of the coefficient is relatively large in all the different regression models, ranging from 0.32 to 0.49. That clearly indicates a strong and significant effect of the geographical distance, even if we control for technological proximity. In other words, knowledge spillovers in China are strongly localized.

In the same line, the coefficient of the technological proximity variable is positive and significant at 1% level in all the specifications. That indicates that, as expected, technological proximity has a positive impact on knowledge flows between Chinese Provinces. When two Provinces share the same technological structure, when they specialize in the same technology, the likelihood that they cite each other is more important.

Table 4 Regression results (1).

	(1) Tobit	(2) Tobit	(3) Negative binomial	(4) Negative binomial	(5) Generalised negative	(6) Generalised negative
	$ln(C_{ij})$	$ln(C_{ij})$	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Co-patenting	0.001	0.001***	0.0004	0.0005*	0.0002	0.0003*
	(1.27)	(2.59)	(1.46)	(1.69)	(1.20)	(1.82)
ln(Pi*Pj)	0.691***	1.955***	0.698***	0.723***	0.701***	0.743***
	(6.66)	(4.26)	(10.10)	(12.14)	(12.13)	(14.88)
Sij	2.470***	0.608***	3.137***	2.885***	3.109***	2.900***
	(3.86)	(7.99)	(6.75)	(6.25)	(7.09)	(6.69)
ln(dij)	-0.490***	-0.435***	-0.388***	-0.386***	-0.316***	-0.369***
	(-3.19)	(-3.79)	(-3.01)	(-3.26)	(-3.72)	(-4.49)
F1i	0.219**		0.189**		0.199***	
	(2.07)		(2.16)		(3.15)	
F2i	0.135		0.158		0.239***	
	(0.98)		(1.57)		(2.99)	
F1j	0.245**		0.239***		0.209***	
	(2.30)		(3.61)		(3.61)	
F2j	0.229		0.237***		0.177**	
	(1.63)		(2.61)		(2.07)	
F1i × F1j		0.176***		0.150***		0.119***
		(3.37)		(3.28)		(3.96)
F2i × F2j		0.130***		0.104***		0.072***
		(3.58)		(3.79)		(4.18)
Traffici	0.130	0.117**	0.110	0.094*	0.142**	0.065
	(1.16)	(1.99)	(1.29)	(1.84)	(2.34)	(1.56)
Trafficj	0.198*	0.130**	0.232***	0.162***	0.153**	0.114***
	(1.68)	(2.16)	(3.00)	(2.75)	(2.29)	(2.64)
Constant	-4.965*	-4.073**	-6.805***	-7.123***	-7.775***	-7.486***
	(-1.94)	(-2.16)	(-3.67)	(-4.05)	(-5.56)	(-5.43)
Observations	930	930	930	930	930	930
Geographical position	Control	Control	Control	Control	Control	Control
lambda	1.300***	0.961***				
	(4.11)	(4.81)				
ln(Pi*Pj)					-0.581***	-0.596***
					(-9.40)	(-8.89)
Sij					2.880**	3.413**
					(2.22)	(2.38)
ln(dij)					0.516**	0.481**
					(2.17)	(2.12)

Robust z-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5 Regression results (2).

	(1) Tobit	(2) Negative binomial	(3) Generalised negative binomial
	$ln(C_{ij})$	C_{ij}	C_{ij}
Co-patenting	0.001	0.001**	0.0003*
	(1.48)	(1.96)	(1.66)
$ln(P_i*P_j)$	0.666***	0.700***	0.688***
	(5.39)	(10.14)	(11.31)
S_{ij}	2.305***	2.981***	2.958***
	(3.33)	(5.88)	(6.25)
$ln(d_{ij})$	-0.431***	-0.309**	-0.257***
	(-2.64)	(-2.39)	(-2.83)
$F1_i$	0.264**	0.246***	0.261***
	(2.27)	(2.79)	(3.97)
$F2_i$	0.106	0.152	0.235***
	(0.71)	(1.49)	(2.97)
$F1_j$	0.295**	0.295***	0.268***
	(2.50)	(4.36)	(4.97)
$F2_j$	0.211	0.215**	0.177**
	(1.40)	(2.43)	(2.43)
D1	-0.077	-0.178	-0.044
	(-0.27)	(-1.01)	(-0.29)
D2	0.800***	0.813***	0.648***
	(2.71)	(4.30)	(3.56)
D3	0.079	-0.083	-0.107
	(0.36)	(-0.54)	(-0.68)
$Traffic_i$	0.205	0.214**	0.234***
	(1.62)	(2.50)	(3.29)
$Traffic_j$	0.283**	0.336***	0.252***
	(2.09)	(4.24)	(3.62)
Constant	-5.583*	-7.886***	-8.412***
	(-1.91)	(-4.15)	(-5.70)
Observations	930	930	930
Geographical position	Control	Control	Control
lambda	1.374***		
	(4.11)		
$ln(P_i*P_j)$			-0.588***
			(-8.44)
S_{ij}			3.554**
			(2.32)
$ln(d_{ij})$			0.448*
			(1.89)

Robust z-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

More originally, we can see in Table 4 that the research structure of Chinese Provinces do impact the intensity of knowledge flows. In particular, F1 is positive and significant both for the citing and cited side, thus suggesting that Provinces that experience a good performance with regard to public research are more likely to both cite and be cited by other Provinces. The same result, although in a less clear-cut way (F2 is only significant in the case of the Generalized Negative Binomial), can be observed for private research.

Furthermore, an interesting result is given by the interaction terms between F1 i and j and F2 i and j, which are both positive and significant. This suggests that when two regions share the same research infrastructure, for instance when they experience both good performance with regard to public or private research, they are more likely to exchange knowledge. This result can be completed by results from Table 5 where we can see that coefficient of variable D2 is positive and strongly significant while coefficient of D1 is negative but not significant. This suggests that knowledge flows are particularly strong between Chinese Provinces that are good at conducting private research but not necessarily between Provinces that are good at conducting public research. This surprising result (the interaction terms between F1 i and j is positive while D1 is non significant) could be explained by the fact that D1 is a dichotomic variable that contains less information than the interaction term. This shows that citations between 'public' provinces only work when the infrastructure is of a high level (high score of F1 i and j) and that it is not enough that both provinces have a minimum of public infrastructure (positive D1). This result is consistent with the fact that patent and patent citations focus mostly on industrial innovation and less on scientific research. Maybe a different result would be observed by analyzing scientific publications citations. Another possible interpretation is that as public research is more science based than private one (Pavitt, 1984), it is possible that public spillovers go beyond Chinese borders. In other words, Chinese patents cite technological advances from Chinese

Provinces and scientific advances from around the world. Another interesting result concerns the absence of a significant effect of the variable D3, which suggests that the regional combination of a high level of public and private research infrastructures does not guarantee a significant knowledge exchange flow. This surprising result is however consistent with previous findings by Guittelman (2006) suggesting that patent citations are associated with specific combinations of individuals' knowledge and organizational structures. A high level of public and private research institutions in some region is not a sufficient condition of high knowledge flows. Further research is needed to identify the specific organizational conditions (especially between the private and public spheres) that enable a high level of knowledge exchange.

A last result deals with the effect of social proximity. The coefficient of the co-patenting variable is very low thus suggesting a weak effect of prior social links. Furthermore, in contrast with former studies, the introduction of social proximity in the analysis does not affect and diminish the importance of geographical and technological proximity. The coefficients of these two variables are not affected when we introduce social proximity in the regression (see appendix Ⅲ, where the results of the regression without social proximity are displayed). This is a very surprising result compared to previous works, and, in any case, a Chinese specificity. For the moment, we have no interpretation of this phenomenon and further work, both theoretical and empirical, will undoubtedly be needed in order to understand this Chinese originality.

II.5. Conclusion

This chapter investigated the determinants of knowledge spillovers between 31 Chinese administrative Provinces (Hong-Kong, Macao and Taiwan excluded). By looking at USPTO patent citations between years 1975-2019, we have been able to put forward, at the level of Chinese Provinces, strong geographical and technological effects. In addition, the importance of social proximity is not validated by our study, which stands in sharp contrast with other studies. Last, an important and original result of our research is that, in order to understand knowledge flows between regions, not only the technological, geographical and/or sectoral dimension matters, but also the type of research that is conducted in the region. Indeed, by conducting a factor analysis, we have been able to develop an indicator of the public versus private research specialization of each Chinese Province. By doing so, we have been able to show that the Provinces' research structure (more or less strong with regard to public and or private research) strongly and significantly impact knowledge flows between Provinces. In particular, knowledge flows seem to be stronger when the two Provinces rank good with regard to private research. This result, which to our knowledge has never been put forward in the literature, might suggest the importance of the nature of the research that is conducted at the regional level in order to explain knowledge flows between regions.

Even if this research suffers from many limitations (for instance we do not control for citations added by USPTO examiners, we focus only on first inventors in the patent, or we remain at the province level) it contributes to feed discussions on the uneven economic development of China. Indeed, it is well known that localized knowledge spillovers are a major source of economic concentration and therefore of divergence between regions. This effect seems to be strong in China. Clearly, knowledge is not equally distributed in space, and not easily accessible at every point in space. As a consequence, the difference between Province that are leading and Provinces that are lagging behind is still increasing. In the less developed regions of China, governments should consider setting up incentives and mechanisms for fostering local and extra-local knowledge transfer networks to create learning regions. In order to fight against these divergences and to ensure an even smooth development of all Chinese Provinces, Chinese authorities might attempt to significantly improve the knowledge flows from most advanced to less developed regions in order to help the latter breaking out their “locked-in” development trajectories.

In the same vein, the very small effect of social proximity in our analysis might suggest that interpersonal mobility remains too weak between Chinese Provinces. Geographical constraints on knowledge flows can be overcome by fostering interpersonal links across regional boundaries. However, this is possible only if there is a minimum level of social interactions between Provinces. Therefore, Chinese authorities could think about feeding and nurturing inter-regional inter-personal networks through encouraging mobility and interaction of people outside their own regions, for instance through subsidizing joint R&D projects, joint conferences and encouraging local firms to tap into external collaborative networks.

Last, and even if more research needs to be conducted in order to confirm this effect, it seems that the research specialization of Chinese Provinces does impact knowledge flows which seem to be more important between Provinces good at private research. This might suggest the existence of a locking-in of some Province that remain specialized in public research and others specialized in private research. Again, appropriate economic policies could aim at unlocking the situation and favoring interactions and knowledge flows between Provinces specialized in public research and the ones specialized in private research.

Appendix I. Comprehensive scores and ranking of regional innovation (following the factor analysis)

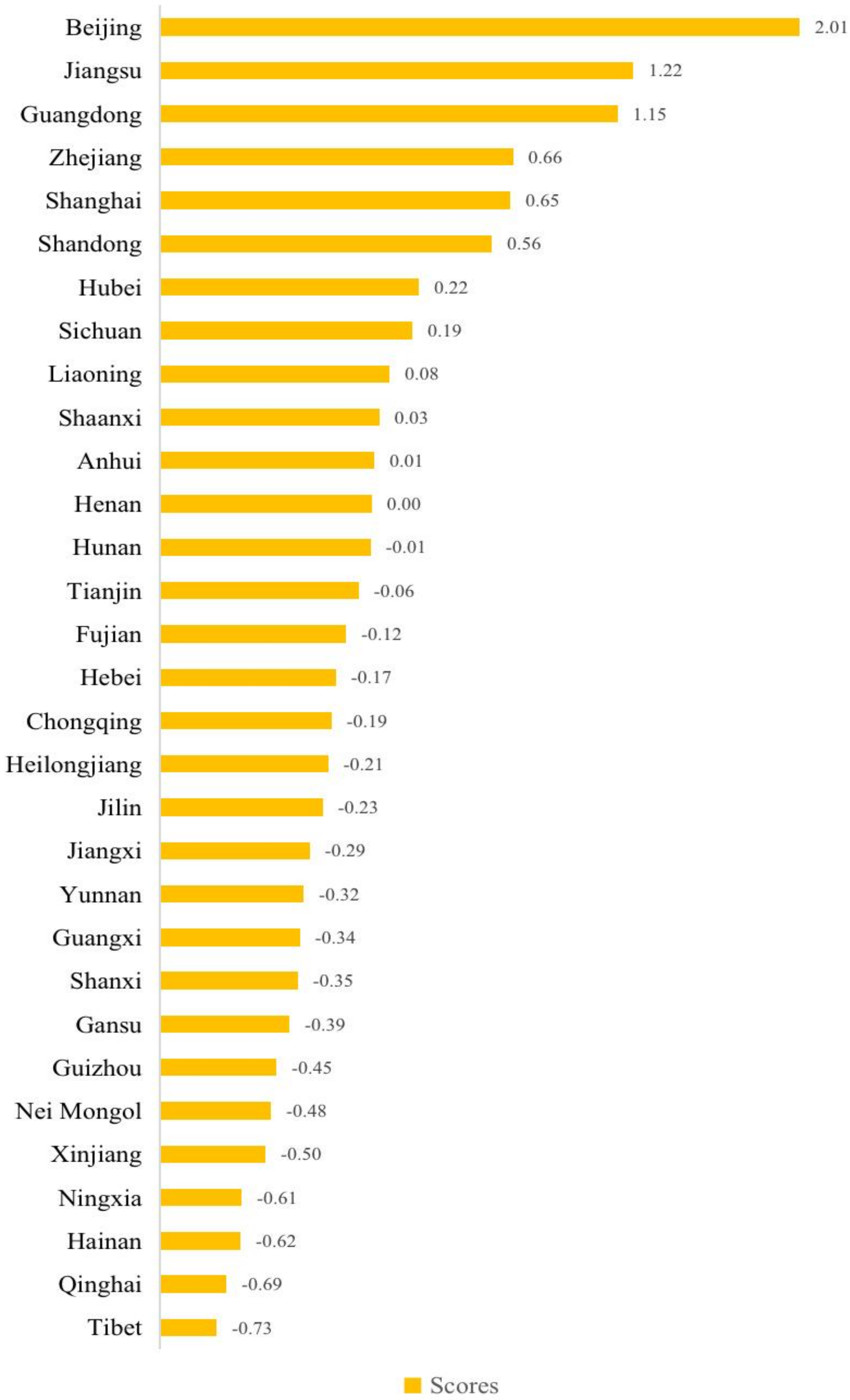

Figure 2. Comprehensive scores and ranking of regional innovation (following the factor analysis)

NOTE: The comprehensive score indicates the relative position of the region to the average level. The higher the comprehensive score, the higher the Province's innovation capability. The function is as follows:
Factor=0.40437Factor1+0.35198Factor2+0.15012Factor3

Appendix II. Some descriptive statistics of variables used in the analysis

Descriptive Statistics on Patent Citations in China.

	N	Mean	Std. dev.	Min	Max.
Total sample of citation links	961	28.22	317.05	0	6439
inter-province citation links	930	7.11	44.61	0	762
Positive inter-province citation links	243	27.22	84.08	0	762
Intra-province citation links	31	661.52	1625.41	0	6439

Degree Centrality and Intra-province citation of the network, 1995–2019.

Region	Abbreviation	Indegree	Outdegree	Total citation	Intra-province citation	Proportion of intra province Citation (%)
Guangdong	G.D.	1744	2263	8183	6439	78.69
Beijing	B.J.	1269	1193	7181	5912	82.33
Shanghai	Sh.H.	1202	1016	4938	3736	75.66
Zhejiang	Zh.J.	647	512	2238	1591	71.09
Jiangsu	J.S.	661	622	2138	1477	69.08
Fujian	F.J.	223	158	779	556	71.37
Hubei	Hu.B.	292	115	411	119	28.95
Sichuan	S.C.	134	279	270	136	50.37
Shandong	Sh.D.	96	148	207	111	53.62
Shaanxi	Sh.X.	53	23	133	80	60.15
Tianjin	T.J.	42	42	123	81	65.85
Liaoning	L.N.	40	26	99	59	59.60
Hebei	He.B	9	24	77	68	88.31
Jiangxi	J.X.	43	10	68	25	36.76
Hunan	Hu.N.	30	32	49	19	38.78
Anhui	A.H.	35	40	48	13	27.08
Chongqing	Ch.Q.	23	27	31	8	25.81
Henan	He.N	13	18	24	11	45.83
Guangxi	G.X.	11	14	23	12	52.17
Shanxi	S.X.	20	7	23	3	13.04
Gansu	G.S.	7	6	15	8	53.33
Heilongjiang	H.L.J.	3	8	13	10	76.92
Jilin	J.L.	3	22	13	10	76.92
Xinjiang	X.J.	2	1	12	10	83.33
Nei Mongol	N.M.G.	4	2	7	3	42.86
Ningxia	N.X.	2	0	7	5	71.43
Guizhou	G.Zh.	3	4	3	0	0.00
Yunnan	Y.N.	2	1	3	1	33.33
Hainan	Hai.N.	1	1	2	1	50.00
Qinghai	Q.H.	0	0	0	0	0.00
Tibet	T.B.	0	0	0	0	0.00

Citation Received and Citation Made Lag Distribution by 4-Year Sub-periods

	Grant Years of Cited-Backward Patents.					
	1996-99	2000-03	2004-07	2008-11	2012-15	2016-19
Lag(years)						
0	7.7	6.7	4.3	5.8	3.6	4.4
1	46.2	20.0	26.3	28.5	21.0	21.7
2	30.8	21.3	26.0	21.9	23.1	21.0
3	15.4	22.7	13.9	17.5	15.8	14.7
4	0.0	18.7	13.2	11.4	11.2	10.4
5	0.0	9.3	9.6	6.7	7.6	8.3
6	0.0	0	3.6	3.2	6.2	6.0
7+	0.0	1.3	3.2	4.9	11.4	13.6
Total	100.0	100.0	100.0	100.0	100.0	100.0
	(ii) Mean and Standard Deviation of the Lag, in Years.					
Mean	1.54	2.61	2.67	2.66	3.37	3.52
s.d.	0.84	1.50	1.77	2.08	2.61	2.87

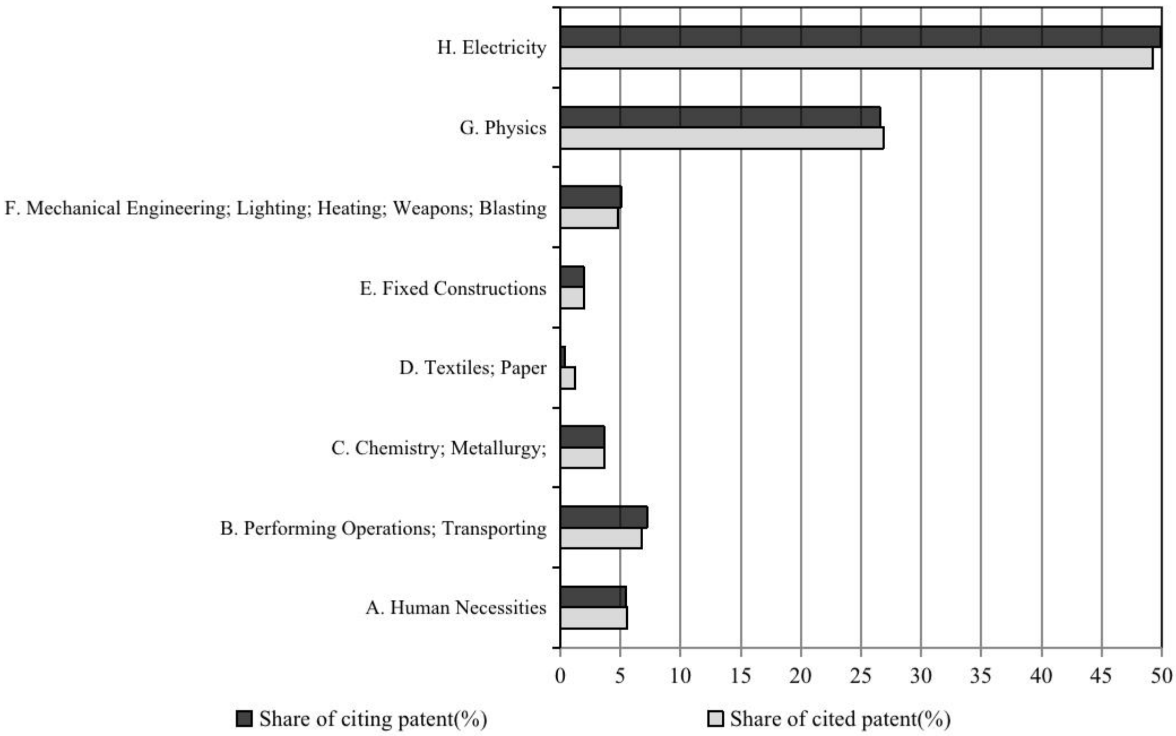

Figure 3. Distribution of citations by technological field (IPC).

Technology Proximity Distribution of Patent Citation in China.

Year	S_{ij}<0.8	0.8<S_{ij}<0.899	S_{ij}>0.9	Sum
1995	0	0	0	0
1996	0	0	3	3
1997	0	0	6	6
1998	2	0	18	20
1999	2	1	19	22
2000	3	1	27	31
2001	3	3	38	44
2002	1	4	67	72
2003	1	7	89	97
2004	3	5	167	175
2005	4	8	474	486
2006	6	31	494	531
2007	8	40	633	681
2008	11	34	947	992
2009	7	51	1218	1276
2010	16	115	1378	1509
2011	12	94	1356	1462
2012	12	150	1347	1509
2013	13	152	1595	1760
2014	24	191	2129	2344
2015	35	284	2613	2932
2016	24	275	3334	3633
2017	16	329	3747	4092
2018	16	151	2281	2448
2019	3	63	927	993
Sum	222	1989	24907	27118

Geographic Distribution of Patent Citation in China.

Year	Citation <1000km	Citation of 1000-1500km	Citation >1500km	Sum
1995	0	0	0	0
1996	3	0	0	3
1997	6	0	0	6
1998	20	0	0	20
1999	22	0	0	22
2000	28	0	3	31
2001	41	1	2	44
2002	62	7	3	72
2003	90	5	2	97
2004	168	3	4	175
2005	469	13	4	486
2006	485	38	8	531
2007	625	41	15	681
2008	930	44	18	992
2009	1183	65	28	1276
2010	1334	132	43	1509
2011	1247	144	71	1462
2012	1247	168	94	1509
2013	1413	228	119	1760
2014	1846	282	216	2344
2015	2310	376	246	2932
2016	2863	459	311	3633
2017	3429	377	286	4092
2018	2067	221	160	2448
2019	860	71	62	993
Sum	22748	2675	1695	27118

Appendix Ⅲ. Regression results without co-inventor as independent variable

Estimation results of Chinese patent citations (Dependent variable: Patent Citation).

	(1) Tobit	(2) Tobit	(3) Negative binomial	(4) Negative binomial	(5) Generalised negative	(6) Generalised negative
	$ln(C_{ij})$	$ln(C_{ij})$	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Co-patenting						
$ln(P_i*P_j)$	0.774***	0.719***	0.712***	0.741***	0.718***	0.771***
	(6.39)	(7.80)	(10.58)	(13.12)	(12.72)	(17.04)
S_{ij}	2.885***	2.326***	3.245***	3.005***	3.243***	3.092***
	(3.49)	(3.73)	(6.82)	(6.29)	(7.34)	(6.94)
$ln(d_{ij})$	-0.551***	-0.492***	-0.375***	-0.375***	-0.287***	-0.331***
	(-2.59)	(-3.02)	(-2.88)	(-3.13)	(-3.58)	(-4.11)
$F1_i$	0.309**		0.201**		0.211***	
	(2.25)		(2.31)		(3.36)	
$F2_i$	0.178		0.162		0.239***	
	(0.94)		(1.59)		(2.91)	
$F1_j$	0.341**		0.258***		0.225***	
	(2.46)		(3.89)		(3.96)	
$F2_j$	0.315*		0.243***		0.178**	
	(1.67)		(2.64)		(2.07)	
$F1_i \times F1_j$		0.239***		0.168***		0.137***
		(3.40)		(3.72)		(4.49)
$F2_i \times F2_j$		0.131**		0.101***		0.066***
		(2.52)		(3.70)		(3.94)
$Traffic_i$	0.184	0.146*	0.120	0.103**	0.154**	0.080**
	(1.22)	(1.76)	(1.42)	(2.04)	(2.53)	(2.02)
$Traffic_j$	0.275*	0.155*	0.241***	0.166***	0.165**	0.126***
	(1.78)	(1.83)	(3.16)	(2.92)	(2.57)	(3.18)
Constant	-6.226*	-5.636**	-7.296***	-7.661***	-8.556***	-8.646***
	(-1.87)	(-2.20)	(-4.01)	(-4.41)	(-7.01)	(-7.04)
Observations	930	930	930	930	930	930
Geographical position	Control	Control	Control	Control	Control	Control
lambda	1.810***	1.382***				
	(4.95)	(4.74)				
$ln(P_i*P_j)$					-0.579***	-0.587***
					(-9.45)	(-9.21)
S_{ij}					2.943**	3.586**
					(2.30)	(2.52)
$ln(d_{ij})$					0.520**	0.510**
					(2.17)	(2.25)

Robust z-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

Estimation results of Chinese patent citations (Dependent variable: Patent Citation).

	(1) Tobit	(2) Negative binomial	(3) Generalised negative binomial
	$ln(C_{ij})$	C_{ij}	C_{ij}
Co-patenting			
$ln(P_i{*}P_j)$	0.770***	0.720***	0.720***
	(5.08)	(10.60)	(12.73)
S_{ij}	2.759***	3.150***	3.219***
	(3.07)	(5.99)	(6.93)
$ln(d_{ij})$	-0.481**	-0.292**	-0.221**
	(-2.16)	(-2.22)	(-2.55)
$F1_i$	0.367**	0.260***	0.272***
	(2.41)	(2.93)	(4.16)
$F2_i$	0.153	0.159	0.236***
	(0.75)	(1.52)	(2.88)
$F1_j$	0.403***	0.318***	0.285***
	(2.61)	(4.57)	(5.22)
$F2_j$	0.302	0.226**	0.183**
	(1.49)	(2.50)	(2.48)
D1	-0.116	-0.123	-0.023
	(-0.29)	(-0.72)	(-0.15)
D2	0.971**	0.783***	0.590***
	(2.41)	(4.23)	(3.34)
D3	-0.0003	-0.109	-0.171
	(-0.00)	(-0.71)	(-1.19)
$Traffic_i$	0.281*	0.224***	0.241***
	(1.65)	(2.58)	(3.28)
$Traffic_j$	0.380**	0.344***	0.259***
	(2.13)	(4.36)	(3.76)
Constant	-7.265*	-8.581***	-9.593***
	(-1.91)	(-4.52)	(-7.60)
Observations	930	930	930
Geographical position	Control	Control	Control
lambda	1.892***		
	(4.94)		
$ln(P_i{*}P_j)$			-0.584***
			(-8.48)
S_{ij}			3.680**
			(2.45)
$ln(d_{ij})$			0.479**
			(2.01)

Robust z-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

Chapter Ⅲ: Knowledge Spillovers Through Co-patenting Among Inventors in China

Ⅲ.1. Introduction

The importance of R&D collaborations for innovation is widely recognized in recent years. R&D collaboration contributes to innovation in the means of sharing resources and expertise (Meli, 2000; Beaver, 2001; Sonnenwald, 2007), exchanging ideas (Melin, 2000; Birnholtz, 2007), learning new skills (Heinze & Kuhlmann, 2008), pooling expertise for complex problems (Sonnenwald, 2007), and facilitating knowledge flows (Singh, 2005; Montobbio and Sterzi, 2011). Existing research has explored R&D cooperation in various contexts. The major current focus in R&D collaboration has been conducted in the aspects of the explanation of the growth (Monjon and Waelbroeck, 2003; Loof and Heshmati, 2002; Belderbos et al., 2015), the measurement of the R&D collaboration (Cronin et al., 2003; Savanur and Srikanth, 2009), the factors that influence R&D collaborations (Amabile et al., 2001; Birnholtz, 2007; Stokols et al., 2008; Bammer, 2008; Rigby, 2009), and the reasons for collaboration (Beaver, 2001; Sonnenwald, 2007; Beaver, 2001). Collaboration with various R&D partners enables companies to search for various sorts of knowledge for innovation. Recent empirical studies have explored the influence of R&D cooperation with different types of collaboration partners, including vertical partners, competitors, and institutional partners (Aschoff and Schmidt 2008; Belderbos, 2015; Franco and Gussoni, 2014).

Patents are the major inventive output of industries, and are easy to commercialize (Griliches, 2007). Existing studies are mostly concerned with R&D collaboration via the publication of co-authored articles (e.g. Wagner 2005; Heinze and Bauer 2007; Mattsson et al. 2008) and sometimes also Community Innovation Survey. However, literature specific inventors' collaboration on patenting is relatively scarce. Most existing research in co-patenting is about the social network connected by the inventors. The major differences between this study and previous studies are as follows: First, we use the patent as units of analysis, in contrast to most previous studies that use aggregate data to study inventor cooperation at the city, region level. Secondly, we consider different types of collaboration together in one model.

The purpose of this chapter is to fill in the gaps in existing research by exploring the extent to which institutional characteristics explain the inventor engagement in collaborative knowledge production, and which organization collaboration patterns are more likely to collaborate across regions. This analysis is based on the patent data applied by China's inventors in the USPTO between 1995 and 2019 to explore the factors that influence the knowledge flows through inventor collaboration in China using logit models. The major findings of this chapter are as follows: First, this chapter shows that universities are more likely to collaborate in patenting, and industries prefer patents independently. Second, inventors affiliated with various universities have a positive and statistically significant influence on the likelihood of interregional R&D collaboration. Third, the inventors from University-Industry and Inter-industries have a significantly negative impact on the probability of interregional R&D collaboration. In other words, inter-industry and UI cooperation are more likely to occur within provincial boundaries. This finding has significant implications for the policy. The results are robust to the robustness checks.

This chapter is structured as follows: Section 2 discusses empirical background on R&D cooperation. Section 3 presents the empirical model and data. Section 4 discusses the empirical results. Section 5 concludes the chapter and discusses the directions for further research.

Ⅲ.2. R&D cooperation theory

Many studies have examined the impact of of R&D cooperation on innovation (e.g., Belderbos, 2004; Becker, 2004; Almeida, 2011; Belderbos, 2015; De Noni, 2017; Petruzzelli, 2020; Audretsch, 2020). However, increasing productivity isn't the only incentive for companies to collaborate. Also, collaboration encourages inter-organizational learning, the generation of new ideas, and, as a result, the development of new innovative products and services (Hardy et al. 2003; Singh and Fleming, 2010). The majority of prior study has examined the frequency of R&D cooperation at different level to investigate which factors are more effective. Muscio (2013) investigates the role of cognitive distance on collaborations. Belderbos (2006) tests whether different types of R&D cooperation are complements in improving the productivity of firms. Jiang（2017）explores regional factors that affect cross-city R&D collaborations in China using co-patent data.

A patent is a good indicator of innovation output since it is externally verified and represents technical advances. Co-patenting can be considered an indicator of collaboration spillovers (Basche, 2021). Inventors who engage in co-patenting can combine their complementary and diversified sources of knowledge (Belderbos, 2014; Boschma, 2005). However, literature specific inventors' collaboration is relatively scarce. The existing literature on patent collaboration is mostly concerned with social networks at the regional or national level. For instance, Yao (2020) explores the impact of extra-local interactions in intercity co-invention networks on cities' innovation performance. Guan (2012) analyzes patent cooperation networks at the country level to capture the profiles of the international knowledge flows.

Prior research demonstrates that knowledge spillovers positively affect growth and overall innovative performance, and cooperative arrangements for technical innovation attracted significant theoretical and empirical attention in recent decades. According to the literature on innovation, R&D cooperation between actors is an important channel of knowledge spillovers (Tether, 2002; Singh, 2005; Montobbio, 2011;).

The literature concentrated mostly on R&D collaboration with various individual partners but ignored the characteristics between partners. Lastly, there exists literature analyzing the determinants of co-inventor tie formation(Cassi et al., 2010).

Geographic proximity is widely mentioned in the literature as a factor that fosters interactions, knowledge creation, and innovation. Existing research suggests that collaboration is closely linked to geographic distance, which is based on the argument that knowledge is partially tacit and localized, and that its transfer requires frequent face-to-face interactions and personal contacts (Kogut & Zander, 1992; Alcacer, 2007). Bathelt (2004) points out that the combination of local interaction and extra-regional interactions can result in a process of knowledge creation and interactive learning that is important to the success of clusters.

After providing an overview of the general motivations for co-patenting with external partners, we will look at the various characteristics of patents and how they may relate to inventors' willingness to engage in co-patenting for innovation.

Ⅲ.3. Data and Method

Ⅲ.3.1 Data

The USPTO dataset is well-suited for the investigation of our research objectives and has several advantages that we exploit in our research. The first advantage is that using USPTO data instead of data from the Chinese patent office allows selecting patents with higher expected economic value. Indeed, since international extensions are costly, Chinese patent applicants should extend their patent in the USA only when they believe the patent is worth it. Furthermore, since USPTO is the world's highest-level patent organization and since the US is China's most important destination country for patent applications overseas, focusing on USPTO data allows us to keep enough patents in our database.

In most cases, there is no formal agreement among the inventors, however, they are frequently involved in the development of a patent. For a patent, if have two or more inventors, there is co-inventor cooperation. Assuming that inventors listed on the same patent are acquainted, co-inventor relationships can be considered as a channel for the transfer of tacit knowledge. The network of co-inventors is an interpersonal network of individuals, who collaborate and exchange information to create new products, and these inventors work in universities, research institutes, or industrial R&D departments. Each patent provides information on the inventors, their name, city, and country. Using the city of residence of the inventors we manually located them in China's provincial administrative units. One major advantage of using patents is that the addresses of inventors are systematically recorded in these texts. Compared with the assignee's location, the address of inventors can better capture the location R&D takes place. The patent also offers information on assignees, for which it was determined whether they were private companies, universities and research institutes, or individuals. According to the patent's International Patent Classification (IPC) codes, we can identify the patent's technological fields.

Given the focus of the chapter, the empirical analysis was carried out at the patent level. The dataset under investigation consists of all utility patents granted by the United States Patent and Trademark Office (USPTO), with all the inventors in the patent reporting a Chinese address. In particular, the actual data set on the patent was downloaded from

PatentView. The data refers to the period of 1995-2020, which covers the majority of patents submitted by China. The number of applications has gone up significantly since 2006, and we can better analyze the trend.

Since we focus on the inventors' cooperation within China, we define "co-inventor" as the cooperator whose country code is "CN" among all the inventors of a patent in this chapter. We don't include the patents, which have inventors from a foreign country. The data elements that we utilize are (1) the patent number; (2) the patent grant year and application year, which can be obtained by patent number; (3) the resident city of each inventor; (4) the International Patent Classification (IPC) codes, and (5) the name of assignee organization.

We record the number of inventors and assignees, the number of IPC6, the detail length, and the reference count it makes to previous patents based on the information disclosed in each patent. All inventors' addresses listed in the patent are uniquely assigned to one out of 31 provincial administrative units in China based on the corresponding city names. In addition, we distinguish the type of assignees and divided them into four categories of industry, university, individual, and overseas institution.

The total number of observations on patents is 93,926. Of them, 66,291 are co-inventor relationships, and 4,789 are cross-region co-inventor relationships. Table 6 provides the number of inventors per patent observation, which ranges from a minimum of one inventor up to a maximum of 35 different inventors. In the patent sample, approximately 29.42% are invented by only one inventor. Also, 66,291 patents (or 70.57% of the sample) are invented by two or more inventors. It is important to emphasize that there is an increasing amount of total China invented patents filed in the US, and the collaborative activity with other inventors increases more quickly than inventing alone (Figure 4). In addition, the amount of intra-provincial co-inventor relationships increases sharply than inter-provincial ones (Figure 5).

Ⅲ.3.2 Model and variables

The empirical analysis explores the factors that impact whether a patent has a co-inventor relationship and whether the co-inventor relationships cross regional borders. As a result, the dependent variables are (1) a binary variable equal to 1 if a patent has two and more inventors residents in China and 0 otherwise; (2) a binary variable equal to 1 if inventors of a patent come from different provincial administrative units in China and 0 otherwise. Given the binary nature of the dependent variable, a logit regression is used to determine the effect of different variables on the likelihood of forming co-inventor connections, and the factors that affect the co-inventor relationships across regional borders.

There are three key independent variables considered to impact the co-inventor relationships: (1) university and industry co-ownership (UI), (2) industry and industry co-ownership (II), and (3) university and university co-ownership (UU).

The literature points out several variables that may affect collaboration. Therefore, we inserted further variables in our model, controlling for patent-related effects. In particular, at the patent level, we consider the technology scope, measured as the number of six-digit IPC codes to which the patent belongs (Moser and Nicholas 2004). The second control is the backward citations of each patent. Moreover, this study controls for the effects of the length of the description text. We also include a control variable related to the size of the team engaged in the creation of the patent (TeamSize), measured as the number of inventors (Mariani 2004). It is also important to control for differences across technology areas when estimating the likelihood that a patent has a co-inventor cooperation. Patents are categorized by highly precise technology classes and therefore may be utilized in measurements of innovative activities in particular economic sectors. We classify patents into 8 eight unique industries as specified by the International Patent Classification (IPC) system. Since a patent may correspond with multiple IPC codes, it is possible for a patent to fall into more than one broad technology classification. As a result, technology classification dummies–eight in all – are included in the regression analysis (Briggs, 2015). Finally, to control for exogenous shocks related to temporal dynamics, year fixed effects, which are associated with the year patent was applied, are also included in all estimation specifications. In the next section, I will present the principal findings of the current study.

Ⅲ.4. Empirical Results

Table 7 provides the summary statistics for the variables included in our model, and Table 8 presents the bivariate correlations between the variables used in the estimation to identify possible multicollinearity problems in the covariates. All correlations are well within the allowed range and can be included in the regression analysis. For each variable in the model, we computed the variance inflation factor (VIF), which was lower than the threshold value of 10, indicating that multi-collinearity does not contaminate the results as suggested by Mason and Perreault (1991).

The first set of questions aimed to analyze what factors affect the probability of China's inventors cooperating. Table 9 presents the results estimating the likelihood a patent is invented by more than two inventors with a logit model. The model main focus on the effect of the institutional characteristics variable, and consider sectoral and time period fixed effect. As for the coefficients of the institutional dummy variables, University has a positive and statistically significant impact (at the 1% level) on the likelihood that a patent has a coinventor relationships. Specifically, inventor from university is found to increase the likelihood that a patent has a co-inventor relationship by 30.7%. In contrast, firms, individuals, and foreign governments have a negative impact. In particular, inventors from firms are found to reduce the likelihood that a patent has co-inventor relationships by 21.5%, inventors from overseas organizations are found to reduce the likelihood that a patent has co-inventor relationships by 14.7%, and inventors working without an organization are found to reduce the likelihood that a patent has co-inventor relationships by 23.2%. It implies that inventors from university prefer to collaborate on patents, and inventors from overseas organizations, or work as individuals are more likely to work alone. As regards the control variables, when technological scope increases by 1 unit, the probability of inventor cooperation increases by 0.4%.

The second set of analyses examined which factors have an impact on cross-region knowledge flow through patent collaborations. Table 10 shows the result of determinants on the formation of cross-regional inventor collaboration. Looking at the coefficients for the co-ownership dummy variables, we find significant positive signs for the UU co-ownership. The UU co-ownership increases the likelihood that a patent has cross-region co-inventor relationships by 6.4%. It indicates that collaboration between universities is more likely to occur across provincial borders. The UI co-ownership and II co-ownership have a significant negative effects on the probability of the cross-region co-inventor relationship, and decrease the probability by 4.2% and 11.8% respectively. One interpretation is that industries are more likely to search for collaboration with university in local regions, and industries rely more on local networks. As regards the control variables, there are also some interesting and conclusive findings. A patent with a higher level of technological complexity is more likely to result in local collaboration, even if the effect is limited in magnitude. By contrast, the more inventors on a patent, the more likely it will involve cross-region collaboration. Additionally, a patent that references a greater number of prior patents is more likely to be an inter-region collaborative work. However, the detail length has no impact on the cross-regional collaboration.

Robustness analysis

The results from Table 9 and Table 10 are robust to a variety of sensitivity checks. In this section, we investigate the robustness of our empirical results. We replace the dependent variable as robustness checks. We change the dependent variable (Dummy variable for the patent inventors, coded 1 if the patent has two and more inventors residents in China, 0 if not.) in model 1 with the total number of inventors of each patent. We run a negative binomial regression model for the first model to deal with count-dependent variables affected by the overdispersion problem. Also, we change the dependent variable (Inter-provincial co-inventor dummy variable) in model 2 with the max distance among all inventors, and run the OLS regression.

Table 11 reports the results of determinant of inventors counts of each patent with negative binomial regressions; the effect estimates are very similar to those in Table 9, suggesting that the results are robust. In further robustness checks, Table 12 shows the main results using the max distance among all inventors as the dependent variable. Results do not change to a large extent with respect to the result in Table 10.

Ⅲ.5. Conclusion

A large body of literature emphasizes that knowledge flows affect importantly firm's ability to learn and innovate. This chapter provides attempts to study mechanisms of knowledge spillovers through inventors' cooperation at a patent level. The number of co-inventor patents has steadily increased over time, consistently remaining at approximately 70% of total patents since 2001.

The results of our logit model highlight these findings: inventor from university has a significantly positive impact on the formation of co-inventor relationships. These results suggest that universities are more likely to share knowledge with partners, which is in accord with the university's external social benefit features. On the other hand, inventors from firms are significantly negatively associated with the formation of co-inventor relationships. In accordance with the present results, previous studies have demonstrated that joint patent ownership is considered the second-best choice to sole ownership for the firms (Kristie, 2015).

The UU co-ownership is shown to positively impact the formation of cross-region co-inventor relationships. This finding suggests that universities are more likely to collaborate on the patents with other universities with relatively long distances. The UI co-ownership and II co-ownership has a significant negative effect on the formation of cross-region co-inventor relationship. These results indicate that industries prefer to seek potential partners within a short distance, regardless of the sort of partner.

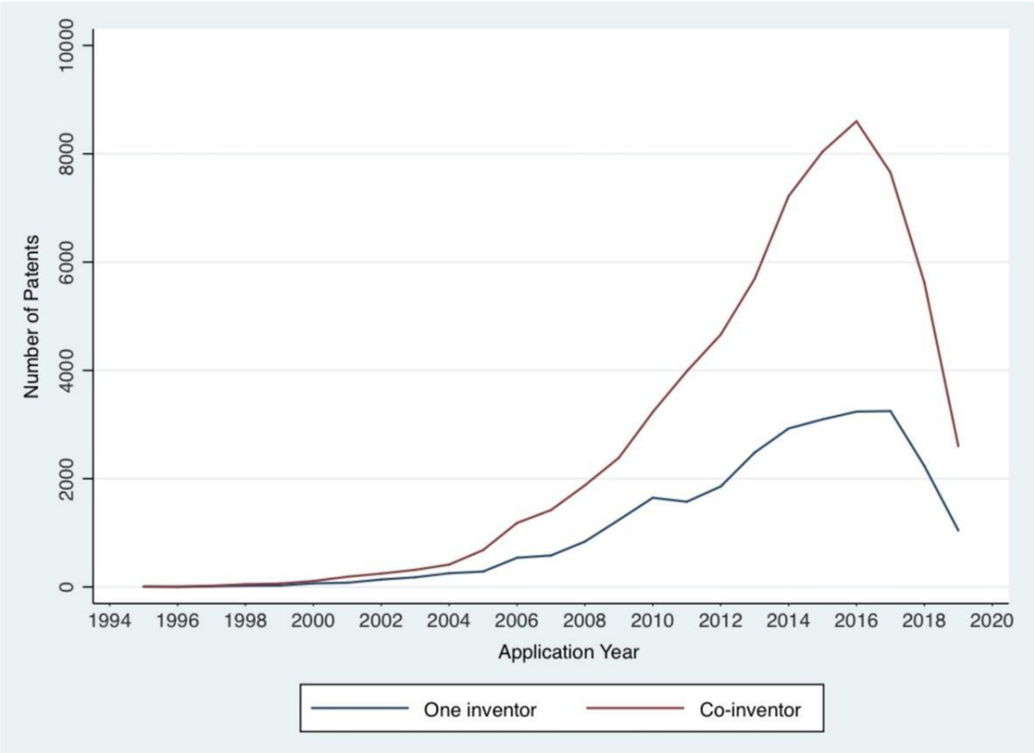

Figure 4. Distribution of patents by co-inventor or not in the period of 1995-2019

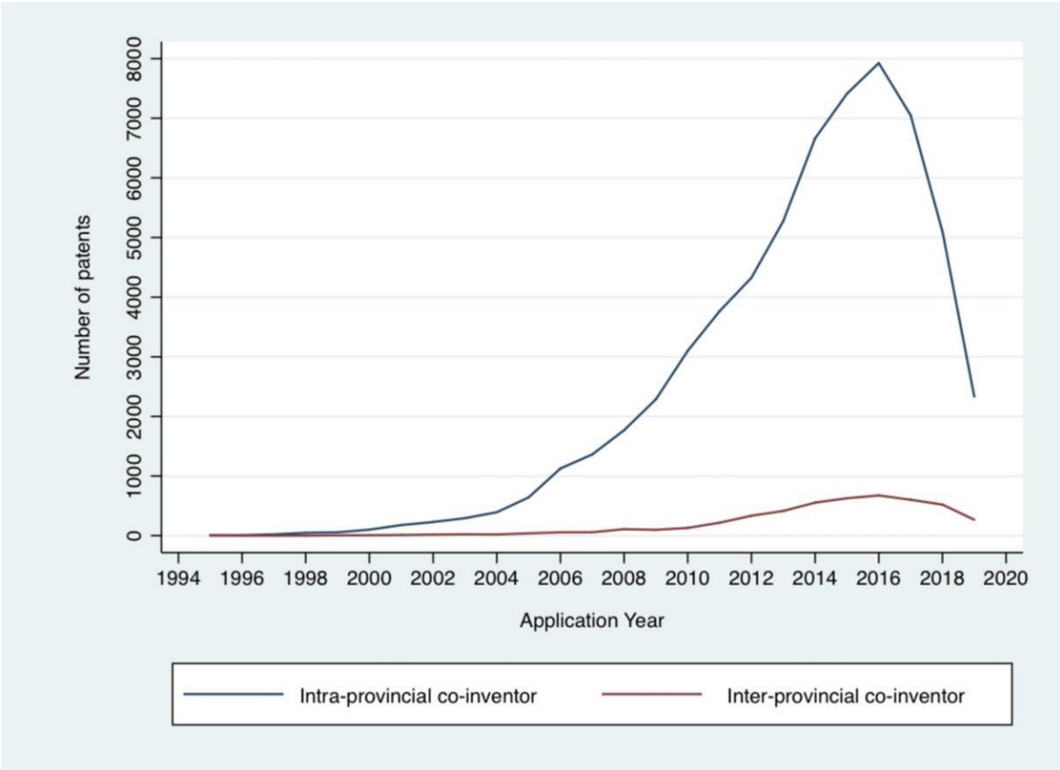

Figure 5. Distribution of patents by Intra-provincial collaboration and Inter-provincial Collaboration in the period of 1995-2019

Table 6 Number of inventors per patent.

Number of inventors	Frequency (observations)	Percent of total observations
1	27635	29.42%
2	19713	20.99%
3	17660	18.80%
4	12044	12.82%
5	6978	7.43%
6	3933	4.19%
7	2080	2.21%
8	1352	1.44%
9	912	0.97%
10	588	0.63%
11	393	0.42%
12	235	0.25%
13	135	0.14%
14	90	0.10%
15	61	0.06%
16	39	0.04%
17	15	0.02%
18	15	0.02%
19	11	0.01%
20	14	0.01%
21	6	0.01%
22	5	0.01%
23	1	<0.01%
24	1	<0.01%
25	1	<0.01%
26	1	<0.01%
28	1	<0.01%
29	1	<0.01%
30	3	<0.01%
31	2	<0.01%
35	1	<0.01%
Total	93926	100%

Table 7 Descriptive statistics of dependent and independent variables（Obs：93,926）

Variable	Definition	Mean	Std dev	Min	Max
Dependent variable					
Co-inventor	Dummy variable for the patent inventors, coded 1 if the patent has two and more inventors residents in China, 0 if not.	0.71	0.46	0	1
Inter-provincial co-inventor	Dummy variable for the patent inventors, coded 1 if the inventors of the patent comes from different provincial administrative units in China, 0 if not.	0.05	0.22	0	1
Inventor count	The number of inventors in each patent.	2.98	2.17	1	35
Max coinventor distance	The max distance among all inventors.	637.9	3,025	0	33,009
Independent variable					
UI	Dummy variable for the patent assignee, coded 1 if there is a university and industry collaboration.	0.03	0.17	0	1
II	Dummy variable for the patent assignee, coded 1 if there is a intra-industry collaboration.	0.10	0.31	0	1
UU	Dummy variable for the patent assignee , coded 1 if there is an intra-university collaboration.	0.002	0.04	0	1
Assignee count	The number of assignees of each patent.	1.15	0.38	1	13
University	Dummy variable for the patent assignee , coded 1 if there is a university in the assignee.	0.10	0.30	0	1
Firms	Dummy variable for the patent assignee , coded 1 if there is a firm in the assignee.	0.73	0.44	0	1
Individual	Dummy variable for the patent assignee , coded 1 if there is an individual in the assignee.	0.01	0.10	0	1
Overseas	Dummy variable for the patent assignee , coded 1 if there is an assignee from foreign.	0.22	0.41	0	1
Technological scope	The number of 6-digit IPC technology fields.	4.01	2.90	1	63
Detail length	Length of the description text.	29,403	29,677	95	1,425,652
Reference count	Citations made to US granted patents by US patents.	4.391	12.64	0	725

Table 8 Correlation matrix of independent variables.

		1	2	3	4	5	6	7	8	9	10
1	UI										
2	II	-0.06									
3	UU	-0.01	-0.01								
4	Assignee count	0.44	0.79	0.10							
5	University	0.54	-0.11	0.13	0.19						
6	Firms	-0.09	0.20	-0.07	0.12	-0.42					
7	Individual	-0.02	-0.03	0.00	0.09	-0.02	-0.12				
8	Overseas	0.11	0.03	-0.02	0.06	-0.05	-0.74	-0.05			
9	Technological scope	0.00	-0.03	0.01	-0.02	0.05	0.04	0.01	-0.11		
10	Detail length	-0.01	-0.06	-0.01	-0.07	-0.06	0.02	-0.03	-0.01	0.20	
11	Citation	0.00	-0.04	-0.01	-0.03	-0.03	-0.10	0.01	0.13	-0.02	0.02

Table 9 Co-inventor dummy, 1995-2019 (patent level)

Variable	Variable Coefficient Marginal Effects	Odds-ratio
Assignee count	0.117***	1.830
	(21.33)	
University	0.307***	4.884
	(22.78)	
Firms	-0.216***	0.328
	(-22.07)	
Individual	-0.232***	0.301
	(-14.70)	
Overseas	-0.147***	0.468
	(-15.83)	
Technological scope	0.004***	1.020
	(6.02)	
Detail length	0.000***	1.000
	(31.54)	
Reference count	-0.0003**	0.9998
	(-2.58)	
Year dummies	Included	
8 sector dummies	Included	
Observations	93,926	

Average marginal effects. z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10 Inter-provincial co-inventor dummy, 1995-2019 (patent level)

	Variable Coefficient Marginal Effects	Odds-ratio
UI	-0.042***	0.519
	(-5.25)	
II	-0.118***	0.161
	(-14.72)	
UU	0.064***	2.679
	(4.90)	
Assignee count	0.050***	2.160
	(8.78)	
Individual	0.032***	1.654
	(3.35)	
Overseas	0.022***	1.403
	(9.33)	
Technological scope	-0.0008**	0.988
	(-2.04)	
Detail length	0.000***	1.000
	(23.20)	
Inventor count	0.004***	1.071
	(9.79)	
Reference count	0.0001***	1.002
	(2.65)	
Year dummies	Included	
8 sector dummies	Included	
Observations	66,291	

z-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 11 Inventor count, 1995-2019 (patent level)

	Negative binomial regression
Assignee count	0.270***
	(42.17)
University	0.258***
	(22.90)
Firms	-0.296***
	(-24.76)
Individual	-0.478***
	(-19.19)
Overseas	-0.280***
	(-24.46)
Technological scope	0.007***
	(8.47)
Detail length	0.000***
	(32.22)
Reference count	-0.000
	(-1.61)
Constant	0.827***
	(5.82)
Year dummies	Included
8 sector dummies	Included
Observations	93,926

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12 Co-inventor max distance, 1995-2019 (patent level)

	(1) OLS
UI	-48.460***
	(-3.84)
II	-131.966***
	(-11.99)
UU	159.812***
	(5.38)
Assignee count	70.342***
	(7.24)
Individual	43.889***
	(2.89)
Overseas	-7.228**
	(-2.12)
Technological scope	-0.731
	(-1.34)
Detail length	0.002***
	(33.42)
Inventor count	5.481***
	(8.00)
Reference count	0.039
	(0.35)
Constant	25.029
	(0.25)
Year dummies	Included
8 sector dummies	Included
R-squared	0.039
Observations	66,291

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Chapter VI: Knowledge Spillovers of Regional Innovation:

The Cases Of China

VI.1. Introduction

In the existing literature on regional technology diffusion, the main determinants of innovation are R&D input, human capital, and the various channels that facilitate knowledge spillovers. Studies on technological spillovers also emphasize that these to a large extent are geographically localized (Jaffe 1989; Krugman, 1991). The most frequently advanced explanation for the regional localization of knowledge is the tacit nature of knowledge, which is acquired through direct, interpersonal contact (Anselin et al., 2000). Given that knowledge spillovers are not directly observable, it is difficult to find systematic evidence of the extent and significance of these impacts.

The objective of this chapter is to shed some further light on the issue in a Chinese context. This work is one of the first attempts to apply spatial economic analysis of R&D knowledge spillovers to developing countries. The contribution of this study to the literature is to examine the extent to which patent citation based spillover and inter-personal connections contribute to innovation production across regions.

This chapter investigates how R&D activities in other areas influence the innovation performance of Provinces through various knowledge spillovers channels. Specifically, patent citation based spillover and inter-personal connections, R&D input from neighboring regions are the main mechanisms we concentrate on. We collect data from 31 Chinese provinces as evidence for our theory. We estimate knowledge spillovers in China by means of a spatial Durbin model and knowledge production function framework. We use patent applications in these provinces from 2006 to 2019 as proxies for knowledge production, patent citation, and inventor collaboration to represent knowledge spillovers. The finding of this chapter shows that knowledge spillovers through patent citation and inventor collaboration have a positive and significant impact on our measure of innovative activities.

We organize the remainder of this chapter as follows. Section 2 introduces the relevant background literature to develop the foundations for an 'extended' RKPF approach. Section 3 describes the data for our study on knowledge spillovers in China, and identifies the most suitable econometric approach for the empirical estimation. Section 4 provides the results for our main model, its extensions, and some alternative specifications. Section 5 concludes with some policy implications.

VI.2. Knowledge Production Function

The empirical model for exploring knowledge spillovers on regional innovative performance is based on the knowledge production functions (KPF) framework initiated by Griliches (1979) and Jaffe (1989). In the Griliches-Jaffe knowledge production function, R&D expenditure is the key innovation input, while patented innovations are the main innovation output. Numerous investigations have shown a statistically significant and positive relationship between knowledge inputs and innovative outputs. As Audretsch (2003) notes, the knowledge production function has been shown to be most reliable at higher degrees of aggregation, such as in countries and regions. The relationship becomes less strong at the microeconomic level of the enterprise, institution, or even line of business.

Based on the knowledge production function developed by Griliches (1979) and Jaffe (1989), many empirical studies extend this framework to study the effects of knowledge spillovers on regional innovation in various countries (Acs et al., 1992; Anselin et al., 1997; Vargra, 1998; Attant-Bernard, 2001; Anselin, 2000; Fritsch, 2002; Fischer and Varga, 2003; Fritsch, 2007; Cabrer-Borras, 2007; Charlot, 2015). Based on the seminar work of Anselin(1988), lots of paper introduce spatial effects to study the effect of knowledge spillovers on innovation production (Acs, 2002; Fischer, 2003; Moreno, 2005; Varga, 2006; Cabrer-Borras, 2007; Usai, 2011; Autant-Bernard, 2011; Paci, 2014).

'There is no reason that knowledge should stop spilling over just because of borders, such as a city limit, state line or national boundary' (Audretsch and Feldman, 2004). The knowledge that is produced in one region may spill over into another, influencing its innovative performance (Moreno et al., 2005).

The purpose of this chapter is to investigate the influence of knowledge spillovers and, in particular, the effects of R & D performed in surrounding regions on the innovative activities of Provinces. In particular, we examine the spillover effects of R & D expenditures, patent citations, and co-inventor patenting using a modified knowledge production function.

VI.3. Data and empirical methodology

VI.3.1. Data and variables

We create a unique dataset at the provincial level using a mix of publicly data and manually collected data, and we build a balanced panel data of 31 Provinces with 434 observations covering 2006–2019. Our dataset draws on two main sources: the United States Patent and Trademark Office (USPTO) and the National Bureau of Statistics of China (NBS). All variables are on a yearly basis.

Following earlier studies, we define our dependent variable "innovation measurement" as the number of patents field at the USPTO. In detail, based on the original data download from the PatentView, we allocate the patents to the provinces according to the place of residence of inventors, and manually compute the patent counts for each Province. Finally, we converted the dependent variable into the log to guarantee normal distribution.

Moreover, the patent applications are allocated to the residence of inventors. The address of the inventor is more likely to identify the source of the innovation than the address of the company. Because, in large companies, the address of the holding company or head office is usually stated, rather than the location of the research lab. If the place of work and home of the inventor is not in the same district, the spatial distribution of innovative production can be affected (Deyle and Grupp, 2005).

Patents as an innovation measure have several limitations. First, not all inventions are patented and not all patents are transferable to new products or services on the market. Second, only product innovation can be patented, while process and organization innovations can be the most relevant forms of innovation (Charlot, 2015). Last, patenting activity varies by sector, and in terms of innovative content, some patents are used in multiple innovations and others are never used (Buesa and Molero, 2006; Griliches, 1990).

Although patents are not perfect, they are the most commonly used indicator of innovation in the literature, and patents are a reasonably precise indicator of innovation activity (Acs et al. 2002). Therefore, in the absence of more detailed regional data on the characteristics of Chinese innovative activities, R&D expenditures and the number of patents may be viewed as credible indicators of invention inputs and outputs

Independent Variables

To analyze the impact of R&D spillovers on regional innovation production through patent citation and patent cooperation, we present the variables used in the empirical analysis below. These variables correspond to China's 31 provincial administrative units from 2006 to 2019.

R&D input. R&D expenditures are a commonly used measure of innovation inputs and key input in the knowledge production function. Data on regional R & D spending are taken from the China Statistical Yearbook. In addition, we included one-year lagged R&D expenditures in the model. Patent applications may rely on previous years' R&D efforts.

Patent citation. The first measure of knowledge flow is patent citations. A patent citation leaves a paper trail and allows the direction of knowledge flows to be traced. We use the aggregate of the citations made by a Province to the other Provinces each year to quantify information flows from external Provinces to the home Province.

We consider the USPTO patents with at least one inventor residing in China to construct the citation pairs. First, we download the patent citation data from the Patents View website by searching the "Inventor Country-CN" field, and information on patents granted during "January 1995-December 2019". Second, since the USPTO only mentions the name, city, and county of all inventors on patents, we have to manually distinguish the province to which patents belong by inventor city. Once the data is cleaned up, we aggregate the number of citations at the Provinces level.

Patent collaboration. Our second measure of knowledge flow is the collaboration of inventors, which involves face-to-face contact. When inventors collaborate on the same innovation, they share ideas about creating an invention, and the flow of knowledge occurs. The co-inventor data comes from PatentView, and we generate co-inventor pairs using USPTO patents with at least two inventors residing in different Provinces in China. We only look at patents that have inventors from different Provinces. This means that intra-regional collaboration is not considered. We use the number of inventor collaborations in the home Province with other Provinces each year to measure knowledge flows from external Provinces to the home Province.

Population (POP). Population is included to reflect regional size differences (Jaffe 1989; Barrio 2005). Table 13 shows the definition and descriptive statistics of the variables.

Table 13 Descriptive statistics of dependent and independent variables（Obs：434）

Variable	Definition	Mean	Std dev	Min	Max
Dependent variable					
Patent sipo	Number of utility patent applications granted by CNIPA for each region and year	5,363	9,211	4	59,742
Patent uspto	Number of utility patent applications granted by USPTO for each region and year	339.1	885.0	0	5,576
Independent variable					
Citation	The total number of patent citation made to other regions in each year.	26.55	66.09	0	425
Co-inventors	The total number number of co-inventors with other regions in each year.	44.93	109.1	0	641
R&D	R&D expenditure in each region and year	3624791	4931554	4,832	30984890
Population	Number of population in each region and year	4,377	2,812	285	12,489

Table 14 Correlation matrix of independent variables.

	lncitation	lncoinventor	lnR&D expenture	lnpopulation
lncitation	1			
lncoinventor	0.880***	1		
lnR&D expenture	0.747***	0.841***	1	
lnpopulation	0.401***	0.497***	0.718***	1

VI.3.2. Econometric approach

We employ panel data with a spatial model to test the contribution of knowledge spillovers through paper trials and interactions of inventors on innovation creation. As Anselin (1998) demonstrated, the focus of spatial econometrics should be on measuring the effects of spillovers. Therefore, the application of a spatial modeling approach allows us to identify the effect of knowledge spillovers among regions. The spatial autoregressive model (SAR), spatial error model (SEM), and spatial Durbin model (SDM) are the most widely used models in the spatial econometrics literature (LeSage and Pace, 2009, Elhorst, 2014).

In this study, we select the spatial Durbin model (SDM) in our empirical analysis, since the SDM model simultaneously includes the spatial lag terms of the dependent variables and independent variables (Elhorst, 2010), it can therefore calculate direct and spillover effects.

In spatial analysis, one of the most important things is to build the weight matrix. The innovation creation of a Province is associated with the characteristics of its neighbors when spatial spillover effects exist, and the definition of a neighborhood depends on a symmetric weight matrix, introduced exogenously in W. In detail, W is an $N \times N$ matrix that describes the spatial dependency between regions. The w_{ij} elements on the main diagonal are set to zero, and the w_{ij} elements show how Province i is spatially connected to Province j. To normalize the external influence on each Province, the weight matrix is standardized so that the sum of values in each row is equal to one.

In the spatial economic literature, various weight matrices such as contiguity, nearest neighbors, and geographical distance are applied to account for connectivity. This research is based on a binary contiguity spatial weight matrix, which assumes that spatial interaction occurs among provinces that share a common boundary. If two Provinces i and j share a common border, such that wij = 1, a change in the variable of Province i will affect the variable of Province j, and it is zero if otherwise. Consequently, the spatial weight matrix W does not change with time.

VI.3.3. The model

To examine the effect of the R&D effort performed in other regions on the production of USPTO patents by Chinese inventors, following LeSage and Pace's (2009) recommendations, we estimate with the spatial Durbin model. We take the natural logarithms on both sides to ensure that the variables are normally distributed, and the estimated coefficients are expressed in terms of elasticity. Finally, by adding a regional index i and a time index t (year), the specific model for a Province is expressed as follows:

$$
\begin{aligned}
\ln PAT_{i,t} = {} & \beta_0 + \beta_1 \ln Citation_{i,t} + \beta_2 \ln Coinventor_{i,t} + \beta_3 lnR\&D_{i,t-1} \\
& + \beta_4 \ln \mathrm{Population}_{i,t} + \beta_5 \sum_{j=1}^{N} W_{ij} \ln Citation_{i,t} \\
& + \beta_6 \sum_{j=1}^{N} W_{ij} \ln Coinventor_{i,t} + \beta_7 \sum_{j=1}^{N} W_{ij}\, lnR\&D_{i,t-1} \\
& + \beta_8 \sum_{j=1}^{N} W_{ij} \ln \mathrm{Population}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}
\end{aligned}
$$

where p is the spatial autoregressive coefficient, W is an $n \times n$ matrix of exogenous row-normalized spatial weights, and the rest of the notations are the same as mentioned earlier. lnPAT is the natural logarithm of the number of granted patents applications to Provinces at time t; $lnR\&D$ is the natural logarithm of R&D expenditures in each provincial administrative unit at time t-1; lnCitation is the natural logarithm of the total number of patent citations made to other 30 Provinces; lnCoinventor is the natural logarithm of the total number of co-inventors with other 30 Provinces; POP is the total number of population in each Provinces; i index the unit of observation; t indexes time(t=1,..16 for the period 2006-2019); and s is a stochastic error term. Please note that we used one-year lagged R&D expenditures to address the endogenous problem caused by reverse causality. We also include time dummies in all specifications to reflect common fluctuations in the overall economy.

We evaluate the SDM model and carry out the relevant test to select the most suitable spatial model. The likelihood ratio (LR) test and the Wald test are used to determine if SDM can be simplified to SAR or SEM, if $\theta = 0$, then, the SDM degrades into the SLM, and if $\theta+\rho\beta = 0$, it degrades into SEM, otherwise it remains an SDM.

Table 15 Moran’s I for regional patents granted (2006–2019).

Year	Moran’s I value	Z-Value	p-Value
2006	0.385	3.589	0.000
2007	0.388	3.602	0.000
2008	0.370	3.454	0.001
2009	0.340	3.199	0.001
2010	0.366	3.428	0.001
2011	0.363	3.384	0.001
2012	0.323	3.027	0.002
2013	0.373	3.464	0.001
2014	0.331	3.106	0.002
2015	0.328	3.071	0.002
2016	0.403	3.704	0.000
2017	0.424	3.875	0.000
2018	0.416	3.802	0.000
2019	0.377	3.473	0.001

VI.4. Empirical results

We use Moran's I (Moran, 1948) to test if there is a spatial dependence (LeSage and Pace, 2009). Table 15 shows the results of Moran’s I test on the dependent variable. The findings reject the null assumption that this variable is not spatially correlated and provide evidence of spatial autocorrelation of innovation activity across provinces. The Moran’s I test detects positive spatial autocorrelation, which means that the spatial adjacent regions share similar characteristics. That is, the closer they are geographically, the more similar values they may have. If spatial dependence is ignored, regression will result in biased estimators. Consequently, this research uses a spatial econometric model for estimating regional knowledge flows in China.

Table 16 and Table 17 present the spatial Durbin model results for 31 Provinces from 2006 to 2019. In this context, the use of panel data enables us to fully exploit the spatial and time features of the data.

Alternative spatial models, such as the spatial lag model (SAR) and the spatial error model (SEM) are employed, but the spatial Durbin model appears to perform better than any other model. In detail, the Akaike information criterion (AIC) and the Bayesian Information Criterion (BIC) can test which spatial model fits best, and the lower AIC and BIC score of the SDM model suggests that the SDM model works better than the SAR model and the SEM model.

According to the LR test and the Wald test results showed in Table 16, the SDM model does not degrade into the SAR model or SEM model, which manifests the suitability and superiority of the SDM model. The spatial Durbin mode is thus considered the most suitable specification. Consequently, the focus of this study is the spatial Durbin model.

Table 16 presents the estimation of the spatial Durbin model of the explanatory variables and the average of these variables from the neighboring Provinces. From columns (1) in Table 16, we find that on innovative activities the estimates for the patent citation, patent collaboration, and R&D input are all positive and significant. The coefficient for the population is negative and significant. Table 16 also points out the presence of local spatial dependency. The average impacts from the neighboring Provinces, the estimates for $W*lnR\&D_{t-1}$ is the only significant variable, indicating a strong correlation between R&D input located nearby and local innovation production. Indeed, it seems that R&D output in a province is positively influenced by the R&D input of neighboring regions. The result is consistent with the hypothesis that patent citation and patent collaboration themselves represent the knowledge spillover effects, therefore it is normal not significant in the variables from the neighbouring Provinces.

Table 17 displays the direct effects, indirect effects and total effects. The direct effect represents the effect in the own region of a unit change in the explanatory variables. The indirect effect represents the effect in the own region of a unit change in the explanatory variables in all other regions through neighboring relation (LeSage and Pace, 2009). Total effects represent the sum of the two effects. To evaluate the signs and magnitudes of our explanatory variables, we use the results given in Table 17.

In Table 17, regarding the direct effects, we find that the coefficients of R&D input, patent citation, inventor collaboration are positive and statistically significant. In detail, the results indicate that knowledge spillovers through patent citation have a significant effect on innovative activities. The coefficient on the patent citation variable is positive statistically significant at the 1% level. We estimate that a 1% increase in the number of patent citations increases innovation production by about 0.187%. The results show a strong positive spillover effect through inventor collaboration: the estimated coefficient is positive statistically significant at the 1% level. We estimate that a 1% increase in the patent collaboration will generate an increase of 0.258% in innovative activities. The coefficient of R&D input is positive statistically significant at the 1% level. We estimate that a 1% increase in the R&D input will generate a 0.224% increase in innovative activity in terms of patenting.

In terms of the indirect effects, we only have significant values on R&D inputs. These results imply that a 1% increase in R&D input in Province i leads to the increase in innovative activities in the other 30 Provinces by 0.199% in total. As for the total effect, we can estimate that 1% increase in the R&D input will generate a 0.4222% increase in innovative activities in toatl. This means that the higher the R&D input in a Province, the more innovation activities in that Province and the neighbouring Provinces.

These coefficients show that external R&D has a significant additional impact on patent production and, in particular, that citations and co-inventorship patterns are relevant channels of knowledge flow. Consequently, it is clear that a substantial amount of knowledge flows through codified documents, such as public patents, and interpersonal links, such as interregional patent collaborations.

Robustness Test

In Column 2 of Table 16 and Table 18, we replace the dependent variable by the Number of utility patent applications granted by China National Intellectual Property Administration (CNIPA) to have a comparision. The purpose of changing the dependent variable data is to see a more accurate effect of the R&D input. Both the data of R&D input and the new dependent variable originate from China Statistical Yearbook, so the estimated coefficient of R&D input is more significant. Since the independent variables of patent citation and inventor collaboration remain the USPTO data, it is normal for the estimated coefficient of inventor collaboration to become insignificant.

Table 19 shows the results for the benchmark model. The first specification is a non-spatial panel model, estimated by two-way linear fixed effects regression, which controls fixed regional effects and time trend. The second estimate is the Pooled OLS regression model, which shows the relationship between the analysed variables regardless of possible unobservable fixed effects. Both model above control the time effects.

Table 16 Determinants of innovation based on the Spatial Durbin Model with spatial and time fixed-effects (2006-2019).

Variable	(1) Patent USPTO	(2) Patent CNIPA
lncitation	0.187***	0.038**
	(7.39)	(2.13)
lncoinventors	0.258***	0.021
	(8.22)	(0.94)
$\ln R\&D_{t-1}$	0.222**	0.500***
	(2.37)	(7.53)
lnpopulation	-1.173**	-0.970***
	(-2.33)	(-2.75)
W*lncitation	0.035	0.133***
	(0.65)	(3.49)
W*lncoinventors	-0.007	-0.032
	(-0.10)	(-0.69)
$W*\ln R\&D_{t-1}$	0.267**	0.157
	(2.25)	(1.60)
W*lnpopulation	1.120	0.455
	(1.16)	(0.67)
Spatial-rho	-0.170**	-0.026
	(-2.17)	(-0.28)
N	434	434
Log-likelihood	-81.3195	71.9914
R-squared	0.274	0.332
LR spatial lag	13.70***	26.85***
LR spatial error	11.66**	28.64***
Wald spatial lag	12.17**	18.89***
Wald spatial error	8.54 *	26.23***

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Table 17 Direct, Indirect effect, Total Effect Estimates of SDM (Data from USPTO)

Variables	(1) Direct effect	(2) Indirect effect	(3) Total effect
lncitation	0.187***	0.003	0.191***
	(7.18)	(0.07)	(3.90)
lncoinventors	0.258***	-0.040	0.218***
	(8.26)	(-0.71)	(3.79)
$lnR\&D_{t-1}$	0.224**	0.199*	0.422***
	(2.49)	(1.91)	(3.02)
lnpopulation	-1.225**	1.108	-0.117
	(-2.39)	(1.24)	(-0.19)

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18 Direct, Indirect effect, Total Effect Estimates of SDM (Data from CNIPA)

Variables	(1) Direct effect	(2) Indirect effect	(3) Total effect
lncitation	0.038**	0.126***	0.164***
	(2.10)	(3.75)	(4.19)
lncoinventors	0.020	-0.029	-0.010
	(0.91)	(-0.67)	(-0.21)
$lnR\&D_{t-1}$	0.506***	0.140*	0.645***
	(8.00)	(1.74)	(5.85)
lnpopulation	-0.978***	0.428	-0.551
	(-2.84)	(0.64)	(-1.08)

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19 Determinants of innovation based on the two-way fixed effects model (2006-2019).

Variable	(1) Two-way fixed effects model	(2) Pooled OLS
lncitation	0.202***	0.361***
	(7.63)	(12.89)
lncoinventors	0.261***	0.483***
	(7.82)	(13.40)
$lnR\&D_{t-1}$	0.181*	0.603***
	(1.84)	(15.28)
lnpopulation	-0.901**	-0.190***
	(-2.12)	(-4.22)
Constant	7.019**	-4.517***
	(2.11)	(-13.88)
Time fixed effect	Control	Control
Observations	434	434
R-squared	0.416	0.956
F	79.99***	527.69***

t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VI.5. Conclusion

Understanding the factors that stimulate innovation activities has been extensively debated in the economic literature. In this respect, however, the role of spatial interactions of knowledge spillovers has been largely overlooked. Thus, this article starts with the theoretical literature and tests the empirical implications of these questions. The objective of this chapter is to shed further light on the issue of knowledge spillovers between spatially neighboring R&D and innovations. This provided a better understanding of the range of spatial externalities between innovation and R&D in the provinces and the surrounding regions.

In this chapter, we estimate knowledge spillovers in China from a temporal and spatial perspective, by means of a spatial perspective in econometrics and knowledge production function framework. The production function approach focuses on the total output of knowledge generation as a function of R&D input and knowledge spillovers. More specifically, this chapter investigates the role of three mechanisms of knowledge spillovers plays determining the creation of patents in China: patent citations-related spillovers, face-to-face contact spillovers based on inventor collaboration, and R&D input from surrounding regions. This research estimates a spatial Durbin model, not only considering point estimates, but also including the direct, indirect, and total effects to provide more accurate results.

The findings of this chapter suggest that knowledge spillovers through patent citations and inventor collaboration, and R&D input are key ingredients in promoting regional innovative activities. These findings confirm the importance of R&D initiatives for enhancing innovation in China and the necessity of encouraging collaboration to improve the effectiveness of such innovation-enhancing activities. Therefore, regions that want to stay competitive internationally should invest in those aspects of the region's knowledge infrastructure that promote regional knowledge spillovers. Moreover, the presence of spatial knowledge flows in the process of innovation can reduce interregional innovation inequalities.

This chapter also has some limitations that merit further research. First, we consider only a small fraction of the innovative activities of China. Second, we do not consider other knowledge flows mechanism of trade and FDI. We can compare the these different channels together in future work.

General Conclusion

This thesis focused on knowledge spillovers, which is defined as the amount of knowledge that can't be appropriated by the economic agent who created it. We focus on the 'geographic dimensions of knowledge spillovers', and apply these theories to the case of the regions of China. The main contribution is providing a comprehensive understanding of the underlying mechanism of knowledge spillover in China in combination with patent citation, inventor cooperation, and R&D input mechanisms.

We first reviewed potential mechanisms of knowledge spillovers, including foreign direct investment, trade, R&D cooperation, entrepreneurship, and labor mobility, and discuss the measurement problem of knowledge spillovers through the knowledge production function and paper trail approach.

The second chapter investigates the determinants of knowledge spillovers between 31 Chinese administrative Provinces (Hong Kong, Macao, and Taiwan excluded). By looking at USPTO patent citations between the years 1995-2019, we have been able to put forward, at the level of Chinese Provinces, strong geographical and technological effects. In addition, the importance of social proximity is not validated by our study, which stands in sharp contrast with other studies. Last, an important and original result of our research is that to understand knowledge flows between regions, not only the technological, geographical, and/or sectoral dimension matters but also the type of research that is conducted in the region. Indeed, by conducting factor analysis, we have been able to develop an indicator of the public versus private research specialization of each Chinese Province. By doing so, we have been able to show that the Provinces' research structure (more or less strong with regard to public and or private research) strongly and significantly impacts knowledge flows between Provinces. In particular, knowledge flows seem to be stronger when the two Provinces rank good with regard to private research. This result, which to our knowledge has never been put forward in the literature, might suggest the importance of the nature of the research that is conducted at the regional level to explain knowledge flows between regions.

The third chapter explores the extent to which institutional characteristics explain the inventor's engagement in collaborative knowledge production, and which organization collaboration patterns are more likely to collaborate across regions. This analysis is based on the patent data applied by China's inventors in the USPTO between 1995 and 2019 to explore the factors that influence the knowledge flows through inventor collaboration in China using logit models. The major findings of this chapter are as follows: First, universities are more likely to collaborate in patenting, and industries prefer patents independently. Second, inventors affiliated with various universities have a positive and statistically significant influence on the likelihood of interregional R&D collaboration. Third, the inventors from University-Industry and Inter-industries have a significantly negative impact on the probability of interregional R&D collaboration. In other words, inter-industry and UI cooperation are more likely to occur within provincial boundaries.

Combined with the result of the second chapter and the third chapter, the fourth chapter estimate knowledge spillovers in China from a temporal and spatial perspective, using a spatial perspective in econometrics and knowledge production function framework. The production function approach focuses on the total output of knowledge generation as a function of R&D input and knowledge spillovers. More specifically, this chapter investigates the role of three mechanisms of knowledge spillovers plays in determining the creation of patents in China: patent citation-related spillovers, face-to-face contact spillovers based on inventor collaboration, and R&D input from surrounding regions. This chapter estimates a spatial Durbin model, not only considering point estimates, but also including the direct, indirect, and total effects to provide more accurate results.

The findings of this chapter suggest that knowledge spillovers through patent citations and inventor collaboration, and R&D input are key ingredients in promoting regional innovative activities. These findings confirm the importance of R&D initiatives for enhancing innovation in China and the necessity of encouraging collaboration to improve the effectiveness of such innovation-enhancing activities. Therefore, regions that want to stay competitive internationally should invest in those aspects of the region's knowledge infrastructure that promote regional knowledge spillovers. Moreover, the presence of spatial knowledge flows in the process of innovation can reduce interregional innovation inequalities.

To summarize, we proposed in this thesis that knowledge spillovers within China plays an important role in regional innovative activities. We also investigate the determinate of the knowledge spillovers via patent citation and R&D collaboration. More specifically, we focus on effect the geography characteristics of the region, and organization characteristics on knowledge flows. This thesis provides a new perspective to observe the knowledge flows within China and its impact on innovative activities. The finding of this thesis may have an important implication for public policy.

References

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The journal of economic history*, *46*(2), 385-406.

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research policy*, *31*(7), 1069-1085.

Acs, Z. J., Audretsch, D. B., Braunerhjelm, P., & Carlsson, B. (2012). Growth and entrepreneurship. *Small Business Economics*, *39*(2), 289-300.

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1992). Real effects of academic research: comment. *The American economic review*, *82*(1), 363-367.

Acs, Z. J., Braunerhjelm, P., Audretsch, D. B., & Carlsson, B. (2009). The knowledge spillover theory of entrepreneurship. *Small Business Economics*, *32*(1), 15-30.

Agrawal, A., Cockburn, I., & McHale, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, *6*(5), 571-591.

Alcacer, J., & Chung, W. (2007). Location strategies and knowledge spillovers. *Management science*, *53*(5), 760-776.

Alcacer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, *88*(4), 774-779.

Alcácer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in US patents: An overview and analysis. *Research policy*, *38*(2), 415-427.

Ali, M., Cantner, U., & Roy, I. (2017). Knowledge spillovers through FDI and trade: the moderating role of quality-adjusted human capital. In *Foundations of Economic Change* (pp. 357-391). Springer.

Almeida, P., Hohberger, J., & Parada, P. (2011). Individual scientific collaborations and firm-level innovation. *Industrial and Corporate Change*, *20*(6), 1571-1599.

Almeida, P., & Kogut, B. (1997). The exploration of technological diversity and geographic localization in innovation: Start-up firms in the semiconductor industry. *Small Business Economics*, *9*(1), 21-31.

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management science*, *45*(7), 905-917.

Amabile, T. M., Patterson, C., Mueller, J., Wojcik, T., Odomirok, P. W., Marsh, M., & Kramer, S. J. (2001). Academic-practitioner collaboration in management research: A case of cross-profession collaboration. *Academy of management journal*, *44*(2), 418-431.

Anselin, L., & Bera, A. K. (1998). Introduction to spatial econometrics. *Handbook of applied economic statistics*, *237*(5).

Anselin, L., & Griffith, D. A. (1988). Do spatial effects really matter in regression analysis? *Papers in Regional Science*, *65*(1), 11-34.

Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of urban economics*, *42*(3), 422-448.

Anselin, L., Varga, A., & Acs, Z. (2000a). Geographical spillovers and university research: A spatial econometricperspective. *Growth and change*, *31*(4), 501-515.

Anselin, L., Varga, A., & Acs, Z. J. (2000b). Geographic and sectoral characteristics of academic knowledge externalities. *Papers in Regional Science*, *79*(4), 435-443.

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

Atallah, G. (2002). Vertical R&D spillovers, cooperation, market structure, and innovation. *Economics of innovation and new technology*, *11*(3), 179-209.

Audretsch, B. D., & Belitski, M. (2022). The limits to open innovation and its impact on innovation performance. *Technovation*, 102519.

Audretsch, D. B. (2003). Innovation and spatial externalities. *International Regional Science Review*, *26*(2), 167-174.

Audretsch, D. B., & Belitski, M. (2020). The role of R&D and knowledge spillovers in innovation and productivity. *European Economic Review*, *123*, 103391.

Audretsch, D. B., Belitski, M., & Caiazza, R. (2021). Start-ups, innovation and knowledge spillovers. *The Journal of Technology Transfer*, *46*(6), 1995-2016.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, *86*(3), 630-640.

Audretsch, D. B., & Feldman, M. P. (1998). *Innovation in cities: science-based diversity, specialization and localized competition*.

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In *Handbook of regional and urban economics* (Vol. 4, pp. 2713-2739). Elsevier.

Audretsch, D. B., & Fritsch, M. (2002). Growth regimes over time and space. *Regional studies*, *36*(2), 113-124.

Audretsch, D. B., & Lehmann, E. E. (2005). Does the knowledge spillover theory of entrepreneurship hold for regions? *Research policy*, *34*(8), 1191-1202.

Autant-Bernard, C. (2001a). The geography of knowledge spillovers and technological proximity. *Economics of innovation and new technology*, *10*(4), 237-254.

Autant-Bernard, C. (2001b). Science and knowledge flows: evidence from the French case. *Research policy*, *30*(7), 1069-1078.

Autant-Bernard, C., & LeSage, J. P. (2011). Quantifying knowledge spillovers using spatial econometric models. *Journal of regional Science*, *51*(3), 471-496.

Bacchiocchi, E., & Montobbio, F. (2009). Knowledge diffusion from university and public research. A comparison between US, Japan and Europe using patent citations. *The*

Journal of Technology Transfer, *34*(2), 169-181.

Bade, F.-J., & Nerlinger, E. A. (2000). The spatial distribution of new technology-based firms: Empirical results for West-Germany. *Papers in Regional Science*, *79*(2), 155-176.

Balsvik, R. (2011). Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *The Review of Economics and Statistics*, *93*(1), 285-297.

Bammer, G. (2008). Enhancing research collaborations: Three key management challenges. *Research policy*, *37*(5), 875-887.

Barrio-Castro, T. D., & García-Quevedo, J. (2005). Effects of university research on the geography of innovation. *Regional studies*, *39*(9), 1217-1229.

Barrios, S., Dimelis, S., Louri, H., & Strobl, E. (2004). Efficiency spillovers from foreign direct investment in the EU periphery: A comparative study of Greece, Ireland, and Spain. *Review of World Economics*, *140*(4), 688-705.

Barrios, S., & Strobl, E. (2002). Foreign direct investment and productivity spillovers: Evidence from the Spanish experience. *Weltwirtschaftliches Archiv*, *138*(3), 459-481.

Barro, R. J. (1996). Determinants of economic growth: A cross-country empirical study. In: National Bureau of Economic Research Cambridge, Mass., USA.

Basche, H. (2021). Determinants of cross-border co-patents: empirical evidence from 45 European regions. *Review of Regional Research*, 1-22.

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in human geography*, *28*(1), 31-56.

Bator, F. M. (1958). The anatomy of market failure. *The quarterly journal of economics*, 72(3), 351-379.

Baumol, W. J. (2004). Welfare Economics and the Theory of the State. In *The encyclopedia of public choice* (pp. 937-940). Springer.

Beaver, D. (2001). Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, *52*(3), 365-377.

Becker, W., & Dietz, J. (2004). R&D cooperation and innovation activities of firms—evidence for the German manufacturing industry. *Research policy*, *33*(2), 209-223.

Békés, G., Kleinert, J., & Toubal, F. (2009). Spillovers from multinationals to heterogeneous domestic firms: Evidence from Hungary. *World Economy*, *32*(10), 1408-1433.

Belderbos, R., Carree, M., & Lokshin, B. (2004). Cooperative R&D and firm performance. *Research policy*, *33*(10), 1477-1492.

Belderbos, R., Carree, M., & Lokshin, B. (2006). Complementarity in R&D cooperation strategies. *Review of Industrial Organization*, *28*(4), 401-426.

Belderbos, R., Carree, M., Lokshin, B., & Fernández Sastre, J. (2015). Inter-temporal patterns of R&D collaboration and innovative performance. *The Journal of Technology Transfer*, *40*(1), 123-137.

Belderbos, R., Cassiman, B., Faems, D., Leten, B., & Van Looy, B. (2014). Co-ownership of intellectual property: Exploring the value-appropriation and value-creation implications of co-patenting with different partners. *Research policy*, *43*(5), 841-852.

Belenzon, S., & Schankerman, M. (2013). Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and statistics*, *95*(3), 884-903.

Birnholtz, J. P. (2007). When do researchers collaborate? Toward a model of collaboration propensity. *Journal of the American Society for Information Science and Technology*, *58*(14), 2226-2239.

Bjerregaard, T. (2010). Industry and academia in convergence: Micro-institutional dimensions of R&D collaboration. *Technovation*, *30*(2), 100-108.

Blake, A., Deng, Z., & Falvey, R. (2009). How does the productivity of foreign direct investment spill over to local firms in Chinese manufacturing? *Journal of Chinese*

Economic and Business Studies, *7*(2), 183-197.

Blalock, G., & Gertler, P. J. (2008). Welfare gains from foreign direct investment through technology transfer to local suppliers. *Journal of international Economics*, *74*(2), 402-421.

Blalock, G., & Gertler, P. J. (2009). How firm capabilities affect who benefits from foreign technology. *Journal of Development Economics*, *90*(2), 192-199.

Blomström, M. (1986). Foreign investment and productive efficiency: the case of Mexico. *The Journal of Industrial Economics*, 97-110.

Blomström, M., & Kokko, A. (1998). Multinational corporations and spillovers. *Journal of Economic surveys*, *12*(3), 247-277.

Blomstrom, M., Lipsey, R. E., & Zejan, M. (1994). What explains the growth of developing countries? *Convergence of productivity: Cross-national studies and historical evidence*, 243-259.

Blomström, M., & Persson, H. (1983). Foreign investment and spillover efficiency in an underdeveloped economy: Evidence from the Mexican manufacturing industry. *World Development*, *11*(6), 493-501.

Bogers, M., Chesbrough, H., & Moedas, C. (2018). Open innovation: Research, practices, and policies. *California management review*, *60*(2), 5-16.

Boschma, R. (2004). Proximité et innovation. *Économie rurale*, *280*(1), 8-24.

Boschma, R., & Frenken, K. (2010). The spatial evolution of innovation networks: a proximity perspective. In *The handbook of evolutionary economic geography*. Edward Elgar Publishing.

Boschma, R. A. (2005). Does geographical proximity favour innovation? *Économie et institutions*(6-7), 111-128.

Bottazzi, L., & Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review*, *47*(4), 687-710.

Branstetter, L., & Sakakibara, M. (1998). Japanese research consortia: a microeconometric analysis of industrial policy. *The Journal of Industrial Economics*, *46*(2), 207-233.

Branstetter, L. G. (2001). Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the US and Japan. *Journal of international Economics*, *53*(1), 53-79.

Braunerhjelm, P., Acs, Z. J., Audretsch, D. B., & Carlsson, B. (2010). The missing link: knowledge diffusion and entrepreneurship in endogenous growth. *Small Business Economics*, *34*(2), 105-125.

Breschi, S., & Lissoni, F. (2003). *Mobility and social networks: Localised knowledge spillovers revisited*. Citeseer.

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research policy*, *32*(1), 69-87.

Briggs, K. (2015). Co-owner relationships conducive to high quality joint patents. *Research policy*, *44*(8), 1566-1573.

Buesa, M., Heijs, J., & Baumert, T. (2010). The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach. *Research policy*, *39*(6), 722-735.

Buesa, M., Heijs, J., Pellitero, M. M., & Baumert, T. (2006). Regional systems of innovation and the knowledge production function: the Spanish case. *Technovation*, *26*(4), 463-472.

Caballero, R. J., & Jaffe, A. B. (1993). How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. *NBER macroeconomics annual*, *8*, 15-74.

Cabrer-Borras, B., & Serrano-Domingo, G. (2007). Innovation and R&D spillover effects in Spanish regions: A spatial approach. *Research policy*, *36*(9), 1357-1371.

Callejón, M., & Segarra, A. (1999). Business dynamics and efficiency in industries and regions: The case of Spain. *Small Business Economics*, *13*(4), 253-271.

Camagni, R. (1993). Inter-firm industrial networks: the costs and benefits of cooperative behaviour. *Journal of Industry Studies*, *1*(1), 1-15.

Caniëls, M. C. (2000). *Knowledge spillovers and economic growth: regional growth differentials across Europe*. Edward Elgar Publishing.

Cassi, L., & Plunket, A. (2010). The determinants of co-inventor tie formation: proximity and network dynamics.

Caves, R. E. (1974). Multinational firms, competition, and productivity in host-country markets. *Economica*, *41*(162), 176-193.

Charlot, S., Crescenzi, R., & Musolesi, A. (2015). Econometric modelling of the regional knowledge production function in Europe. *Journal of Economic Geography*, *15*(6), 1227-1259.

Chen, X., Liu, Z., & Ma, C. (2017). Chinese innovation-driving factors: regional structure, innovation effect, and economic development—empirical research based on panel data. *The annals of regional science*, *59*(1), 43-68.

Coe, D. T., & Helpman, E. (1995). International r&d spillovers. *European Economic Review*, *39*(5), 859-887.

Coe, D. T., Helpman, E., & Hoffmaister, A. W. (1997). North-south R & D spillovers. *The Economic Journal*, *107*(440), 134-149.

Coe, D. T., Helpman, E., & Hoffmaister, A. W. (2009). International R&D spillovers and institutions. *European Economic Review*, *53*(7), 723-741.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: the two faces of R & D. *The Economic Journal*, *99*(397), 569-596.

Connolly, M. (2003). The dual nature of trade: measuring its impact on imitation and growth.

Journal of Development Economics, *72*(1), 31-55.

Corsino, M., Mariani, M., & Torrisi, S. (2019). Firm strategic behavior and the measurement of knowledge flows with patent citations. *Strategic management journal*, *40*(7), 1040-1069.

Crescenzi, R., Nathan, M., & Rodríguez-Pose, A. (2016). Do inventors talk to strangers? On proximity and collaborative knowledge creation. *Research policy*, *45*(1), 177-194.

Criscuolo, P., Narula, R., & Verspagen, B. (2005). Role of home and host country innovation systems in R&D internationalisation: a patent citation analysis. *Economics of innovation and new technology*, *14*(5), 417-433.

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research policy*, *37*(10), 1892-1908.

Cronin, B., Shaw, D., & La Barre, K. (2003). A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, *54*(9), 855-871.

Dachs, B., Ebersberger, B., & Pyka, A. (2008). Why do firms cooperate for innovation? A comparison of Austrian and Finnish CIS3 results. *International Journal of Foresight and Innovation Policy*, *4*(3-4), 200-229.

Damijan, J. P., Rojec, M., Majcen, B., & Knell, M. (2013). Impact of firm heterogeneity on direct and spillover effects of FDI: Micro-evidence from ten transition countries. *Journal of comparative economics*, *41*(3), 895-922.

De Faria, P., Lima, F., & Santos, R. (2010). Cooperation in innovation activities: The importance of partners. *Research policy*, *39*(8), 1082-1092.

De Loecker, J. (2007). Do exports generate higher productivity? Evidence from Slovenia. *Journal of international Economics*, *73*(1), 69-98.

De Noni, I., Ganzaroli, A., & Orsi, L. (2017). The impact of intra-and inter-regional knowledge collaboration and technological variety on the knowledge productivity of European regions. *Technological forecasting and social change*, *117*, 108-118.

Deyle, H.-G., & Grupp, H. (2005). Commuters and the regional assignment of innovative activities: A methodological patent study of German districts. *Research policy*, *34*(2), 221-234.

Domar, E. D. (1946). Capital expansion, rate of growth, and employment. *Econometrica, Journal of the Econometric Society*, 137-147.

Dunning, J. H. (1993). Internationalizing Porter's diamond. *MIR: Management International Review*, 7-15.

Dunning, J. H. (1998). Location and the multinational enterprise: a neglected factor? *Journal of international business studies*, *29*(1), 45-66.

Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. *Spatial economic analysis*, *5*(1), 9-28.

Elhorst, J. P. (2014). Spatial panel data models. In *Spatial econometrics* (pp. 37-93). Springer.

Fallah, M. H., & Ibrahim, S. (2004). Knowledge spillover and innovation in technological clusters. Proceedings, IAMOT 2004 Conference, Washington, DC,

Fallick, B., Fleischman, C. A., & Rebitzer, J. B. (2006). Job-hopping in Silicon Valley: some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics*, *88*(3), 472-481.

Falvey, R., Foster, N., & Greenaway, D. (2002). North-South trade, knowledge spillovers and growth. *Journal of Economic Integration*, 650-670.

Feldman, M. P. (1994). *The geography of innovation* (Vol. 2). Springer Science & Business Media.

Feldman, M. P. (2001). The entrepreneurial event revisited: firm formation in a regional context. *Industrial and Corporate Change*, *10*(4), 861-891.

Filatotchev, I., Liu, X., Lu, J., & Wright, M. (2011). Knowledge spillovers through human mobility across national borders: Evidence from Zhongguancun Science Park in China. *Research policy*, *40*(3), 453-462.

Fischer, M. M., & Varga, A. (2003). Spatial knowledge spillovers and university research: Evidence from Austria. *The annals of regional science*, *37*(2), 303-322.

Flôres Jr, R. G., Fontoura, M. P., & Santos, R. G. (2000). Foreign direct investment spillovers: what can we learn from Portuguese data?

Fracasso, A., & Marzetti, G. V. (2015). International trade and R&D spillovers. *Journal of international Economics*, *96*(1), 138-149.

Frenken, K., Ponds, R., & Van Oort, F. (2010). The citation impact of research collaboration in science-based industries: A spatial-institutional analysis. *Papers in Regional Science*, *89*(2), 351-271.

Fritsch, M. (2002). Measuring the quality of regional innovation systems: A knowledge production function approach. *International Regional Science Review*, *25*(1), 86-101.

Fritsch, M., & Aamoucke, R. (2013). Regional public research, higher education, and innovative start-ups: An empirical investigation. *Small Business Economics*, *41*(4), 865-885.

Fritsch, M., & Franke, G. (2004). Innovation, regional knowledge spillovers and R&D cooperation. *Research policy*, *33*(2), 245-255.

Fritsch, M., & Lukas, R. (2001). Who cooperates on R&D? *Research policy*, *30*(2), 297-312.

Fritsch, M., & Slavtchev, V. (2007). Universities and innovation in space. *Industry and innovation*, *14*(2), 201-218.

Gao, X., Guan, J., & Rousseau, R. (2011). Mapping collaborative knowledge production in China using patent co-inventorships. *Scientometrics*, *88*(2), 343-362.

Gilbert, B. A., McDougall, P. P., & Audretsch, D. B. (2008). Clusters, knowledge spillovers and new venture performance: An empirical examination. *Journal of Business Venturing*, *23*(4), 405-422.

Gittelman, M. (2006). National institutions, public–private knowledge flows, and innovation performance: A comparative study of the biotechnology industry in the US and France. *Research policy*, *35*(7), 1052-1068.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of political economy*, *100*(6), 1126-1152.

Globerman, S. (1979). Foreign direct investment and'spillover'efficiency benefits in Canadian manufacturing industries. *Canadian journal of economics*, 42-56.

Gomes-Casseres, B., Hagedoorn, J., & Jaffe, A. B. (2006). Do alliances promote knowledge flows? *Journal of Financial Economics*, *80*(1), 5-33.

González-Pernía, J. L., Peña-Legazkue, I., & Vendrell-Herrero, F. (2012). Innovation, entrepreneurial activity and competitiveness at a sub-national level. *Small Business Economics*, *39*(3), 561-574.

Görg, H., & Strobl, E. (2005). Spillovers from foreign firms through worker mobility: An empirical investigation. *The Scandinavian journal of economics*, *107*(4), 693-709.

Gorodnichenko, Y., Svejnar, J., & Terrell, K. (2014). When does FDI have positive spillovers? Evidence from 17 transition market economies. *Journal of comparative economics*, *42*(4), 954-969.

Greunz, L. (2003). Geographically and technologically mediated knowledge spillovers between European regions. *The annals of regional science*, *37*(4), 657-680.

Griffith, R., Lee, S., & Van Reenen, J. (2011). Is distance dying at last? Falling home bias in fixed-effects models of patent citations. *Quantitative economics*, *2*(2), 211-249.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change.

Econometrica, Journal of the Econometric Society, 501-522.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The bell journal of economics*, 92-116.

Griliches, Z. (1990). Hedonic price indexes and the measurement of capital and productivity: some historical reflections. Fifty years of economic measurement: The jubilee of the conference on research in income and wealth,

Griliches, Z. (1991). The search for R&D spillovers. *National Bureau of Economic Research Working Paper Series*(w3768).

Griliches, Z. (2007). R&D and productivity. In *R&D and Productivity*. University of Chicago Press.

Grillitsch, M., & Trippl, M. (2014). Combining knowledge from different sources, channels and geographical scales. *European Planning Studies*, *22*(11), 2305-2325.

Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European Economic Review*, *35*(2-3), 517-526.

Guan, J., & Chen, Z. (2012). Patent collaboration and international knowledge flow. *Information Processing & Management*, *48*(1), 170-181.

Guan, J. C., Yam, R. C., & Mok, C. K. (2005). Collaboration between industry and research institutes/universities on industrial innovation in Beijing, China. *Technology Analysis & Strategic Management*, *17*(3), 339-353.

Haddad, M., & Harrison, A. (1994). Are there Dynamic Externalities from Direct Foreign Investment? Evidence for Morocco. *Market Structure and Industrial Performance*, *15*, 273.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2000). Market value and patent citations: A first look. In: national bureau of economic research Cambridge, Mass., USA.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file:

Lessons, insights and methodological tools. In: National Bureau of Economic Research Cambridge, Mass., USA.

Halpern, L., & Muraközy, B. (2007). Does distance matter in spillover? 1. *Economics of Transition*, *15*(4), 781-805.

Hansen, T., & Hansen, U. E. (2020). How many firms benefit from a window of opportunity? Knowledge spillovers, industry characteristics, and catching up in the Chinese biomass power plant industry. *Industrial and Corporate Change*, *29*(5), 1211-1232.

Hardy, C., Phillips, N., & Lawrence, T. B. (2003). Resources, knowledge and influence: The organizational effects of interorganizational collaboration. *Journal of management studies*, *40*(2), 321-347.

Harhoff, D. (1999). Firm formation and regional spillovers-evidence from Germany. *Economics of innovation and new technology*, *8*(1-2), 27-55.

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economics and statistics*, *81*(3), 511-515.

Harrod, R. (1939). An Essay in Dynamic Theory. *The Economic Journal*, *49*(193), 14-33.

Hartley, J., Sørensen, E., & Torfing, J. (2013). Collaborative innovation: A viable alternative to market competition and organizational entrepreneurship. *Public administration review*, *73*(6), 821-830.

Haskel, J. E., & Slaughter, M. J. (2002). Does the sector bias of skill-biased technical change explain changing skill premia? *European Economic Review*, *46*(10), 1757-1783.

Hausman, J. A., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship. In: national bureau of economic research Cambridge, Mass., USA.

Havranek, T., & Irsova, Z. (2011). Estimating vertical spillovers from FDI: Why results vary and what the true effect is. *Journal of international Economics*, *85*(2), 234-244.

Heinze, T., & Bauer, G. (2007). Characterizing creative scientists in nano-S&T: Productivity, multidisciplinarity, and network brokerage in a longitudinal perspective. *Scientometrics*, *70*(3), 811-830.

Henderson, R., Jaffe, A. B., & Trajtenberg, M. (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. *Review of Economics and statistics*, *80*(1), 119-127.

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of political economy*, *103*(5), 1067-1090.

Hermes, N., & Lensink, R. (2003). Foreign direct investment, financial development and economic growth. *The Journal of Development Studies*, *40*(1), 142-163.

Hoekman, J., Frenken, K., & Van Oort, F. (2009). The geography of collaborative knowledge production in Europe. *The annals of regional science*, *43*(3), 721-738.

Hu, A. G., & Jaffe, A. B. (2003). Patent citations and international knowledge flow: the cases of Korea and Taiwan. *International journal of industrial organization*, *21*(6), 849-880.

Iršová, Z., & Havránek, T. (2013). Determinants of horizontal spillovers from FDI: Evidence from a large meta-analysis. *World Development*, *42*, 1-15.

Jacobs, J. (1969). *The economy of cities* (039470584X).

Jaffe, A. B. (1986). Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value. In: national bureau of economic research Cambridge, Mass., USA.

Jaffe, A. B. (1989). Real effects of academic research. *The American economic review*, 957-970.

Jaffe, A. B., & De Rassenfosse, G. (2019). Patent citation data in social science research: Overview and best practices. *Research handbook on the economics of intellectual property law*.

Jaffe, A. B., & Trajtenberg, M. (1996). Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *proceedings of the National Academy of Sciences*, *93*(23), 12671-12677.

Jaffe, A. B., & Trajtenberg, M. (1999). International knowledge flows: Evidence from patent citations. *Economics of innovation and new technology*, *8*(1-2), 105-136.

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. (2000). Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American economic review*, *90*(2), 215-218.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, *108*(3), 577-598.

Javorcik, B. S., Saggi, K., & Spatareanu, M. (2004). Does it Matter where you come from? Vertical Spillovers from FDI and Investor Nationality. *World Bank Policy Research Working Paper*, *3449*.

Jellema, J., & Roland, G. (2011). Institutional clusters and economic performance. *Journal of Economic Behavior & Organization*, *79*(1-2), 108-132.

Jiang, S.-q., Shi, A.-n., Peng, Z.-h., & Li, X. (2017). Major factors affecting cross-city R&D collaborations in China: evidence from cross-sectional co-patent data between 224 cities. *Scientometrics*, *111*(3), 1251-1266.

Jordaan, J. A. (2005). Determinants of FDI-induced externalities: New empirical evidence for Mexican manufacturing industries. *World Development*, *33*(12), 2103-2118.

Kaiser, U., Kongsted, H. C., & Rønde, T. (2015). Does the mobility of R&D labor increase innovation? *Journal of Economic Behavior & Organization*, *110*, 91-105.

Keller, W. (2010). International trade, foreign direct investment, and technology spillovers. In

Handbook of the Economics of Innovation (Vol. 2, pp. 793-829). Elsevier.

Kim, J., & Marschke, G. (2005). Labor mobility of scientists, technological diffusion, and the firm's patenting decision. *RAND Journal of Economics*, 298-317.

Kokko, A. (1994). Technology, market characteristics, and spillovers. *Journal of Development Economics*, *43*(2), 279-293.

Kokko, A., Tansini, R., & Zejan, M. C. (1996). Local technological capability and productivity spillovers from FDI in the Uruguayan manufacturing sector. *The Journal of Development Studies*, *32*(4), 602-611.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, *99*(3), 483-499.

Lampe, R. (2012). Strategic citation. *Review of Economics and statistics*, *94*(1), 320-333.

Lenaerts, K., & Merlevede, B. (2015). Firm size and spillover effects from foreign direct investment: the case of Romania. *Small Business Economics*, *45*(3), 595-611.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Liu, Z., & Lin, P. (2004). Backward linkages of foreign direct investment: Evidence from China. *Department of Economics, Lingnan University, Hong Kong*.

Lööf, H. (2009). Multinational enterprises and innovation: firm level evidence on spillover via R&D collaboration. *Journal of evolutionary economics*, *19*(1), 41-71.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, *22*(1), 3-42.

MacDougall, G. D. A. (1960). The benefits and costs of private investment from abroad: A theoretical approach 1. *Bulletin of the Oxford University Institute of Economics & Statistics*, *22*(3), 189-211.

MacGarvie, M. (2005). The determinants of international knowledge diffusion as measured by patent citations. *Economics Letters*, *87*(1), 121-126.

Madsen, J. B. (2007). Technology spillover through trade and TFP convergence: 135 years of evidence for the OECD countries. *Journal of international Economics*, *72*(2), 464-480.

Malecki, E. J. (1997). Technology and economic development: the dynamics of local, regional, and national change. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Maliranta, M., Mohnen, P., & Rouvinen, P. (2009). Is inter-firm labor mobility a channel of knowledge spillovers? Evidence from a linked employer–employee panel. *Industrial and Corporate Change*, *18*(6), 1161-1191.

Marshall, A. (2009). *Principles of economics: unabridged eighth edition*. Cosimo, Inc.

Mattsson, P., Laget, P., Nilsson, A., & Sundberg, C.-J. (2008). Intra-EU vs. extra-EU scientific co-publication patterns in EU. *Scientometrics*, *75*(3), 555-574.

Maurseth, P. B., & Verspagen, B. (2002). Knowledge spillovers in Europe: a patent citations analysis. *Scandinavian Journal of Economics*, *104*(4), 531-545.

Medda, G., Piga, C., & Siegel, D. S. (2006). Assessing the returns to collaborative research: Firm-level evidence from Italy. *Economics of innovation and new technology*, *15*(1), 37-50.

Mehta, A., Rysman, M., & Simcoe, T. (2010). Identifying the age profile of patent citations: New estimates of knowledge diffusion. *Journal of Applied Econometrics*, *25*(7), 1179-1204.

Merlevede, B., Schoors, K., & Spatareanu, M. (2014). FDI spillovers and time since foreign entry. *World Development*, *56*, 108-126.

Messeni Petruzzelli, A., & Murgia, G. (2020). University–Industry collaborations and international knowledge spillovers: a joint-patent investigation. *The Journal of Technology Transfer*, *45*(4), 958-983.

Monjon, S., & Waelbroeck, P. (2003). Assessing spillovers from universities to firms: evidence from French firm-level data. *International journal of industrial organization*, *21*(9), 1255-1270.

Montobbio, F., & Sterzi, V. (2011). Inventing together: exploring the nature of international knowledge spillovers in Latin America. In *Catching up, spillovers and Innovation Networks in a Schumpeterian Perspective* (pp. 81-117). Springer.

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, *10*(2), 243-251.

Moreno, R., Paci, R., & Usai, S. (2005). Geographical and sectoral clusters of innovation in Europe. *The annals of regional science*, *39*(4), 715-739.

Morrison, A. (2008). Gatekeepers of knowledge within industrial districts: who they are, how they interact. *Regional studies*, *42*(6), 817-835.

Moser, P., Ohmstedt, J., & Rhode, P. W. (2018). Patent citations—an analysis of quality differences and citing practices in hybrid corn. *Management science*, *64*(4), 1926-1940.

Motohashi, K. (2005). University–industry collaborations in Japan: The role of new technology-based firms in transforming the National Innovation System. *Research policy*, *34*(5), 583-594.

Murata, Y., Nakajima, R., Okamoto, R., & Tamura, R. (2014). Localized knowledge spillovers and patent citations: A distance-based approach. *Review of Economics and statistics*, *96*(5), 967-985.

Nadiri, M. I. (1993). Innovations and technological spillovers. In: National Bureau of Economic Research Cambridge, Mass., USA.

Narin, F., Noma, E., & Perry, R. (1987). Patents as indicators of corporate technological strength. *Research policy*, *16*(2-4), 143-155.

Newman, C., Rand, J., Talbot, T., & Tarp, F. (2015). Technology transfers, foreign investment and productivity spillovers. *European Economic Review*, *76*, 168-187.

Nieto, M. J., & Santamaría, L. (2007). The importance of diverse collaborative networks for the novelty of product innovation. *Technovation*, *27*(6-7), 367-377.

Orlando, M. J. (2004). Measuring spillovers from industrial R&D: on the importance of geographic and technological proximity. *RAND Journal of Economics*, 777-786.

Orlic, E., Hashi, I., & Hisarciklilar, M. (2018). Cross sectoral FDI spillovers and their impact on manufacturing productivity. *International Business Review*, *27*(4), 777-796.

Ott, H., & Rondé, P. (2019). Inside the regional innovation system black box: Evidence from French data. *Papers in Regional Science*, *98*(5), 1993-2026.

Paci, R., Marrocu, E., & Usai, S. (2014). The complementary effects of proximity dimensions on knowledge spillovers. *Spatial economic analysis*, *9*(1), 9-30.

Paci, R., & Usai, S. (1999). Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal*, *49*(4), 381-390.

Park, Y., Yoon, B., & Lee, S. (2005). The idiosyncrasy and dynamism of technological innovation across industries: patent citation analysis. *Technology in Society*, *27*(4), 471-485.

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

Pigou, A. C., & Aslanbeigui, N. (2017). *The economics of welfare*. Routledge.

Plummer, L. A., & Acs, Z. J. (2014). Localized competition in the knowledge spillover theory of entrepreneurship. *Journal of Business Venturing*, *29*(1), 121-136.

Ponds, R., Oort, F. v., & Frenken, K. (2009). Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography*, *10*(2), 231-255.

Porter, M. E. (1990). The competitive advonioge of notions. *Harvard business review*, *73*, 91.

Qiu, S., Liu, X., & Gao, T. (2017). Do emerging countries prefer local knowledge or distant knowledge? Spillover effect of university collaborations on local firms. *Research policy*, *46*(7), 1299-1311.

Rigby, J. (2009). Comparing the scientific quality achieved by funding instruments for single grant holders and for collaborative networks within a research system: Some observations. *Scientometrics*, *78*(1), 145-164.

Roach, M., & Cohen, W. M. (2013). Lens or prism? Patent citations as a measure of knowledge flows from public research. *Management science*, *59*(2), 504-525.

Roberts, M., & Setterfield, M. (2007). What is endogenous growth theory? *Economic growth: new directions in theory and policy*, 14.

Robertson, P. L., & Langlois, R. N. (1995). Innovation, networks, and vertical integration. *Research policy*, *24*(4), 543-562.

Rodríguez-Pose, A., & Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Regional studies*, *42*(1), 51-67.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, *94*(5), 1002-1037.

Romer, P. M. (1990). Endogenous technological change. *Journal of political economy*, *98*(5, Part 2), S71-S102.

Ronde, P., & Hussler, C. (2005). Innovation in regions: What does really matter? *Research policy*, *34*(8), 1150-1172.

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

Santoro, F., Borges, M., & Santos, N. (2000). Cooperation model for learning: A system of patterns. EdMedia+ Innovate Learning, Savanur, K., & Srikanth, R. (2010). Modified collaborative coefficient: A new measure for quantifying the degree of research collaboration. *Scientometrics*, *84*(2), 365-371.

Saxenian, A., & Hsu, J. Y. (2001). The Silicon Valley–Hsinchu connection: technical communities and industrial upgrading. *Industrial and Corporate Change*, *10*(4), 893-920.

Schmiedeberg, C. (2008). Complementarities of innovation activities: An empirical analysis of the German manufacturing sector. *Research policy*, *37*(9), 1492-1503.

Schumpeter, J., & Backhaus, U. (2003). The theory of economic development. In *Joseph Alois Schumpeter* (pp. 61-116). Springer.

Sheshinski, E. (1967). Optimal accumulation with learning by doing. *Essays on the theory of optimal economic growth*.

Sinani, E., & Meyer, K. E. (2004). Spillovers of technology transfer from FDI: the case of Estonia. *Journal of comparative economics*, *32*(3), 445-466.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management science*, *51*(5), 756-770.

Singh, J. (2008). Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Research policy*, *37*(1), 77-96.

Smarzynska Javorcik, B. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American economic review*, *94*(3), 605-627.

Solow, R. M. (1956). A contribution to the theory of economic growth. *the Quarterly journal of Economics*, *70*(1), 65-94.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 312-320.

Song, J., Almeida, P., & Wu, G. (2003). Learning–by–hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management science*, *49*(4), 351-365.

Sonn, J. W., & Storper, M. (2008). The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975–1997. *Environment and Planning a*, *40*(5), 1020-1039.

Sonnenwald, D. H. (2007). Scientific collaboration. *Annu. Rev. Inf. Sci. Technol.*, *41*(1), 643-681.

Steurs, G. (1994). Spillovers and Cooperation in Research and Development, doctoral dissertation KULeuven.

Stiglitz, J. E. (1997). Dumping on free trade: the US import trade laws. *Southern Economic Journal*, *64*(2), 402-424.

Stokols, D., Misra, S., Moser, R. P., Hall, K. L., & Taylor, B. K. (2008). The ecology of team science: understanding contextual influences on transdisciplinary collaboration. *American journal of preventive medicine*, *35*(2), S96-S115.

Stolpe, M. (2002). Determinants of knowledge diffusion as evidenced in patent data: the case of liquid crystal display technology. *Research policy*, *31*(7), 1181-1198.

Stoyanov, A., & Zubanov, N. (2012). Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics*, *4*(2), 168-198.

Swan, T. W. (1956). Economic growth and capital accumulation. *Economic record*, *32*(2), 334-361.

Tether, B. S. (2002). Who co-operates for innovation, and why: An empirical analysis. *Research policy*, *31*(6), 947-967.

Thompson, P. (2006). Patent citations and the geography of knowledge spillovers: Evidence from inventor-and examiner-added citations. *The Review of Economics and Statistics*, *88*(2), 383-388.

Thompson, P., & Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American economic review*, *95*(1), 450-460.

Tijssen, R. J. (2001). Global and domestic utilization of industrial relevant science: patent citation analysis of science–technology interactions and knowledge flows. *Research policy*, *30*(1), 35-54.

Todo, Y., & Miyamoto, K. (2006). Knowledge spillovers from foreign direct investment and the role of local R&D activities: Evidence from Indonesia. *Economic Development and Cultural Change*, *55*(1), 173-200.

Trajtenberg, M. (2001). Innovation in Israel 1968–1997: a comparative analysis using patent data. *Research policy*, *30*(3), 363-389.

Usai, S. (2011). The geography of inventive activity in OECD regions. *Regional studies*, *45*(6), 711-731.

Uzawa, H. (1965). Optimum technical change in an aggregative model of economic growth. *International economic review*, *6*(1), 18-31.

Van Beers, C., Berghäll, E., & Poot, T. (2008). R&D internationalization, R&D collaboration and public knowledge institutions in small economies: Evidence from Finland and the Netherlands. *Research policy*, *37*(2), 294-308.

Van der Panne, G. (2004). Agglomeration externalities: Marshall versus jacobs. *Journal of evolutionary economics*, *14*(5), 593-604.

Varga, A. (1998). *University research and regional innovation: a spatial econometric analysis of academic technology transfers* (Vol. 13). Springer Science & Business Media.

Varga, A. (2006). The spatial dimension of innovation and growth: empirical research methodology and policy analysis. *European Planning Studies*, *14*(9), 1171-1186.

Verspagen, B. (1991). A new empirical approach to catching up or falling behind. *Structural change and economic dynamics*, *2*(2), 359-380.

Veugelers, R., & Cassiman, B. (2005). R&D cooperation between firms and universities. Some empirical evidence from Belgian manufacturing. *International journal of* *industrial organization*, *23*(5-6), 355-379.

Wagner, C. S., & Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research policy*, *34*(10), 1608-1618.

Wu, C.-Y., & Mathews, J. A. (2012). Knowledge flows in the solar photovoltaic industry: Insights from patenting by Taiwan, Korea, and China. *Research policy*, *41*(3), 524-540.

Xiang, X.-Y., Cai, H., Lam, S., & Pei, Y.-L. (2013). International knowledge spillover through co-inventors: An empirical study using Chinese assignees' patent data. *Technological forecasting and social change*, *80*(1), 161-174.

Yao, L., Li, J., & Li, J. (2020). Urban innovation and intercity patent collaboration: A network analysis of China’s national innovation system. *Technological forecasting and social change*, *160*, 120185.

Yasar, M., & Paul, C. J. M. (2007). International linkages and productivity at the plant level: Foreign direct investment, exports, imports and licensing. *Journal of international Economics*, *71*(2), 373-388.

Yu, F., & Wu, Y. (2014). Patent citations and knowledge spillovers: an analysis of Chinese patents registered in the USA. *Asian Journal of Technology Innovation*, *22*(1), 86-99.

Zhang, G., & Tang, C. (2017). How could firm's internal R&D collaboration bring more innovation? *Technological forecasting and social change*, *125*, 299-308.

Zucker, L. G., Darby, M. R., & Armstrong, J. (1998). Geographically localized knowledge: spillovers or markets? *Economic Inquiry*, *36*(1), 65-86.

Université de Strasbourg

Jia LIU

ÉCOLE DOCTORALE AUGUSTIN COURNOT

Knowledge Flow Within China and its Economic Consequences

Résumé

Cette thèse analyse les flux de connaissances par le canal des citations de brevets et de la collaboration des inventeurs sur les brevets dans le contexte de la Chine, puis estime leur contribution aux activités innovantes régionales.

Le premier chapitre donne un aperçu de la littérature existante sur la théorie des externalités de connaissances. Le chapitre 2 analyse les déterminants des flux de connaissances dans le contexte de la Chine.

L'objectif du chapitre 3 est de combler les lacunes des recherches existantes en examinant dans quelle mesure les caractéristiques institutionnelles expliquent l'engagement des inventeurs dans la production collaborative de connaissances, et quels sont les modèles de collaboration entre organisations qui sont plus susceptibles de collaborer entre régions.

Le Chapitre 4 étudie comment les activités de R&D dans d'autres domaines influencent les performances d'innovation des provinces par le biais de divers canaux de diffusion des connaissances.

Mots-clés : diffusion des connaissances ; géographie économique ; R&D structure ; co-inventeur ; citation de brevets

Résumé en anglais

This thesis analyzes knowledge flows through the channel of patent citations and inventors' patent collaboration in the context of China, and then estimates their contribution to regional innovative activities.

The first chapter provides an overview of the existing literature on the theory of knowledge externalities. Chapter 2 analyzes the determinants of knowledge flows in the context of China.

The objective of Chapter 3 is to fill in the gaps in existing research by examining the extent to which institutional characteristics explain inventors' engagement in collaborative knowledge production, and which patterns of collaboration between organizations are more likely to collaborate across regions.

Chapter 4 explores how R&D activities in other fields influence the innovation performance of provinces through various knowledge diffusion channels.

Keywords: knowledge spillovers ; geography economics ; R&D structure ; co-inventor ; patent citation