

Spatial Dependence and Spatial Heterogeneity in the Analysis of Regional Economic Performance and House Price Developments

Inaugural-Dissertation zur Erlangung des akademischen Grades einer
Doktorin der Wirtschaftswissenschaft des Fachbereichs
Wirtschaftswissenschaft der Freien Universität Berlin



vorgelegt von
Katharina Pijnenburg
aus Königs Wusterhausen
2013

Erstgutachter: Prof. Dr. Dieter Nautz
Professur für Ökonometrie

Zweitgutachter: Prof. Dr. Alexander S. Kritikos
Professur für Industrie- und Institutionenökonomie

Tag der Disputation: 14.05.2013

Acknowledgements

I wrote this thesis as a participant of the DIW Graduate Center, and while working as a research assistant in the Macroeconomics Department at DIW Berlin. Here I express my gratitude for all people having supported me in accomplishing my thesis.

First and foremost, I must thank my supervisors Dieter Nautz and Alexander Kritikos for their support, inspiration, and neverending encouragement. In particular, I am thankful for constructive comments and discussions of my research during the seminars of my first supervisor, Dieter Nautz. His feedback strongly encouraged me and stimulated my research. In addition, I am thankful to Alexander Kritikos, who demonstrated me how fruitful coauthorship can be.

Further, I am grateful for the supportive research environment at DIW Berlin and for the diverse opportunities that the Graduate Center offers its participants. In this regard, many thanks also due to Georg Meran, the first Dean of the Graduate Center, for kindly advising me at the beginning of my thesis.

Finally, I thank my parents for supporting my studies and for always being on the spot taking care of my daughter, Marlene, so that I could spend more time doing research. I thank my sister, Franziska, for always believing in me, and my husband, Sebastian, for his patience and for understanding my work related mood swings.

Contents

Acknowledgements	III
List of Figures	VIII
List of Tables	IX
List of Abbreviations	X
Overview	XI
Zusammenfassung	XVI
1 Do Regions with Entrepreneurial Neighbors Perform Better? A Spatial Econometric Approach for German Regions	1
1.1 Introduction	1
1.2 The Spatial Dimension of Entrepreneurship Capital	6
1.3 Model	9
1.4 Spatial Econometric Modeling	11
1.4.1 Spatial Durbin Model	11
1.4.2 Estimation Method	12
1.4.3 Coefficient Interpretation	14
1.4.4 Spatial Weight Matrix Comparison	14
1.5 Estimation Results	19
1.5.1 Coefficient Comparison	19
1.5.2 Model Comparison	21
1.5.3 Results of the Final Model	22
1.6 Conclusion	27
1.A Appendix	29
1.A.1 Figures	29
1.A.2 Tables	31

2	Self-Employment and Economic Performance - A Geographically Weighted Regression Approach for European Regions	32
2.1	Introduction	32
2.2	Necessity and Opportunity Entrepreneurship	37
2.3	Geographically Weighted Regression	39
2.3.1	Description of the Model when Assuming Homogeneous Effects Across Space	39
2.3.2	First Empirical Results when Assuming Homogeneous Effects	40
2.3.3	Description of the Model when Assuming Heterogeneous Effects Across Space	41
2.3.4	GWR Estimation	44
2.3.4.1	Spatial Weighting Function	45
2.3.4.2	Cross-Validation	46
2.3.4.3	Tests on Spatial Heterogeneity	47
2.3.4.4	Empirical Results of the GWR Estimation	48
2.4	The Equilibrium Rate of Entrepreneurship	51
2.4.1	The Model	52
2.4.2	The Estimated Equilibrium Rates of Entrepreneurship	55
2.4.3	Results Using the Equilibrium Rate of Entrepreneurship	56
2.5	Conclusion	60
2.A	Appendix	62
2.A.1	Figures	62
2.A.2	Tables	63
3	The Spatial Dimension of US House Price Developments	65
3.1	Introduction	65
3.2	Theoretical Aspects and Empirical Evidence	68
3.3	Econometric Approach	71
3.3.1	Spatial Fixed Effects Panel Estimation	71
3.3.2	Fixed Effects Panel Smooth Transition Regression Model	73
3.4	Data Description	77
3.5	Empirical Results	80
3.6	Conclusion	87

Contents

3.A Appendix	89
Bibliography	91
List of Publications and Working Papers	XXII
Curriculum Vitae	XXIV
Ehrenwörtliche Erklärung	XXX

List of Figures

1.1	Gross value added per working-age population, 2008	3
1.2	Start-up rates in knowledge intense areas, 1997-2004	7
1.3	The links between regions implied by the best weight matrix . .	23
1.4	Coefficient comparison for different weight matrices: physical capital	29
1.5	Coefficient comparison for different weight matrices: human capital	30
1.6	Coefficient comparison for different weight matrices: entrepreneur- ship capital	30
2.1	Self-employment per working age population, 2004	35
2.2	Plot of GWR parameters of the self-employment rate, α_{ie} , adap- tive Gaussian weighting function	49
2.3	Plot of GWR t-values for the parameters of the self-employment rate, α_{ie} , adaptive Gaussian weighting function	49
2.4	Plot of GWR parameters of the self-employment rate, α_{ie} , adap- tive bi-square weighting function	50
2.5	Plot of GWR t-values for the parameters of the self-employment rate, α_{ie} , adaptive bi-square weighting function	50
2.6	Self-employment rates above and below the constant equilib- rium rate of 12 percent in 2004	58
2.7	Plot of the estimated equilibrium rates of entrepreneurship and the actual rate of self-employment for the year 2004	62
3.1	Annual growth rate of regional house prices, US metropolitan statistical areas	67
3.2	Transition variable - spatially weighted annual real house price growth rate of neighboring regions, $Whpiag_{t-1}$	75
3.3	Transition function vs. transition variable, two extreme regimes	83

List of Figures

3.4	Individual spatial spillover parameters at different points in time, two extreme regimes	86
3.5	Transition function vs. transition variable, three extreme regimes	90

List of Tables

1.1	Coefficient comparison	20
1.2	Posterior model probabilities	22
1.3	Estimation output OLS and of the final Bayesian spatial Durbin model	24
1.4	Data description	31
2.1	Panel estimation	57
2.2	Data description	63
2.3	Individual parameter stationarity test for self-employment . . .	64
3.1	Data description	79
3.2	Spatial panel estimation	81
3.3	PSTR estimation results, two extreme regimes	85

List of Abbreviations

AIC	Akaike Information Criterion
CV	Cross-Validation
FE	Fixed Effects
GWPR	Geographically Weighted Panel Regression
GWR	Geographically Weighted Regression
i.e.	id est
km	Kilometers
MSA	Metropolitan Statistical Area
MUP	Mannheim Enterprise Panel
NLS	Nonlinear Least Squares
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PSTR	Panel Smooth Transition Regression
R&D	Research and Development
SAR	Spatial Lag Model
SARMA	Spatial Autoregressive Moving Average Model
SDM	Spatial Durbin Model
SEM	Spatial Error Model
US	United States

Overview

When analyzing economic aspects at regional levels, the spatial dimension of the data plays a crucial role. This is the case as spatial data and models are often characterized by two spatial effects, namely spatial dependence and spatial heterogeneity. The field of spatial econometrics is a separate field in econometrics that explicitly incorporates these spatial effects into econometric models.

Following Anselin (1988), spatial dependence implies that observations in one region depend on observations in neighboring regions. This is in contrast to the conventional assumption in cross-sectional samples, where it is assumed that observations made in one region are independent of observations in other regions (LeSage and Pace, 2009). Spatial heterogeneity implies that model coefficients vary with location, which is again in contrast to the conventional assumption of homogeneous parameters (Anselin, 1988).

This dissertation consists of three chapters, with each chapter analyzing a different spatial aspect of regional data. The first chapter deals with spatial dependence, with a special focus on the construction and the varying effects of different spatial weight matrices. The second chapter is about spatial heterogeneity. The third chapter combines spatial dependence, spatial heterogeneity, and differences in spatial dependence across time and space into one model. The three chapters analyze different economic aspects but they are linked by the spatial setting.

Chapter 1 analyzes the essential role of entrepreneurship for economic development in German NUTS-3 (Nomenclature of Territorial Units for Statistics) regions. Following Audretsch and Keilbach (2004), a neoclassical production function augmented by entrepreneurship capital is estimated. Entrepreneurship capital should affect economic output, because it facilitates

knowledge spillovers, increases competition, and leads to greater diversity of products and problem solutions. In this chapter the spatial dimension of the data is explicitly taken into account. That means that not only the effect of entrepreneurship capital in a specific region i on economic performance of that region is analyzed, but also the effect of entrepreneurship capital in neighboring regions on economic performance of region i . It is argued that because of the accessibility of entrepreneurship capital and because of competition issues, entrepreneurship capital of neighboring regions may have an effect on economic performance of region i . Theoretical foundations for both positive and negative spatial spillover effects of entrepreneurship capital are provided. A spatial Durbin model is estimated to take the spatial dependence structure of the data into account. Thereby, special emphasis lies on the creation and the choice of the spatial weight matrix. The spatial weight matrix determines to what extent region i affects region j and vice versa. In spatial econometrics, estimates and inference depend on the weight matrix used. The spatial econometric estimation method allows not only to find out how large the effect of entrepreneurship capital is on economic output in a specific region, but also how large the spillover effect is coming from neighboring regions.

Entrepreneurship capital of region i and entrepreneurship capital of neighboring regions are found to have a positive effect on the economic performance of region i . On the one hand, this is interpreted as evidence in favor of knowledge spillover effects via entrepreneurship within one region and between regions. On the other hand, it is argued that improvements in the competitiveness of both region i and regions j will have a positive effect on economic output of region i . Having neighbors with a high level of entrepreneurship capital, i.e., entrepreneurial neighbors, significantly affects a region's performance. However, the significance of the spatial spillover effects largely depends on the choice of the weight matrix. This is seen as evidence that positive and negative spatial spillover effects of entrepreneurship capital cancel out.

The analysis confirms a spatial dependence structure, where a failure to account for it would result in biased estimates.

The contribution of this chapter is that for the first time the spatial dimension of the data in the analysis of entrepreneurship capital on economic performance is explicitly taken into account. Furthermore, the creation of a large number of spatial weight matrices provides evidence to what extent the results depend on the choice of the spatial weight matrix.

This chapter is based on a paper which is joint with Konstantin A. Kholodilin.

Chapter 2 analyzes whether the effect of the self-employment rate on economic performance in European NUTS-2 regions depends on the location. Measures of entrepreneurship, like the self-employment rate, cannot distinguish between necessity and opportunity entrepreneurs. Necessity entrepreneurship does not result in technological change, while opportunity entrepreneurship does (Acs and Varga, 2005). Thus, it is not surprising that most studies analyzing the effect of entrepreneurship on economic development find results that are quite heterogeneous across regions.

Theoretically, a high self-employment rate could mirror a positive entrepreneurial environment but it could also reflect a lack of wage employment opportunities. As long as the positive entrepreneurial environment mainly attracts innovative entrepreneurs, the effect of a high self-employment rate on economic development should be positive. However, in case where a high self-employment rate reflects a lack of wage employment opportunities then the effect on economic development could be insignificant or even negative. This implies that in different regions the effect of the self-employment rate on economic development could be heterogeneous depending on whether the environment attracts innovative entrepreneurs or not.

In this chapter it is analyzed whether the effect of self-employment on economic output in European NUTS-2 regions depends on the location by applying a geographically weighted regression (GWR). Fotheringham et al. (2002) developed GWR to deal with spatial non-stationarity. It is found that regions having a significant positive effect of self-employment on economic development in the GWR estimation have, on average, a lower self-employment rate than regions having a significant negative effect.

In an attempt to find out more about the source of spatial non-stationarity of the effect of self-employment on economic development, subsequently the concept of equilibrium rate of entrepreneurship is used. This concept is applied to estimate a level of the self-employment rate from which point relatively more entrepreneurs are self-employed out of necessity than out of opportunity. It is found that in regions where the self-employment rate is below the equilibrium rate, self-employment is positively associated with economic development. In regions where the self-employment rate is above the equilibrium rate, the effect of self-employment on economic development is negative. This is seen as evidence that self-employment rates above the equilibrium rate can indeed be interpreted as being dominated by entrepreneurs out of necessity. Levels of the self-employment rate below the equilibrium rate imply an environment mainly attracting innovative entrepreneurs.

The contribution of this chapter is that GWR is for the first time applied to visualize the heterogeneous effects of self-employment on economic development on a regional level. Furthermore, this chapter tries to shed light on the question as to why there is spatial heterogeneity by estimating a level of the self-employment rate from which point the rate can be interpreted as being dominated by necessity entrepreneurship.

Chapter 3 is a joint analysis of three spatial characteristics in house price dynamics in US metropolitan statistical areas, namely spatial dependence, spatial heterogeneity, and heterogeneity in spatial dependence. While spatial dependence and spatial heterogeneity are well established aspects of house price developments, differences in spatial dependence across time and space have not attracted much attention yet. It is argued that the disposition effect, which labels the phenomenon in financial markets that investors sell their winning stocks too soon and hold their losing stocks too long, may explain heterogeneous house price spillover across time and space. The disposition effect implies reduced house price spillovers in times of declining house prices. In a first step, a spatial panel regression is estimated to see whether there is overall spatial dependence in house price developments. Subsequently, a

panel smooth transition regression model is applied to estimate heterogeneity in spatial dependence and in the effect of the fundamentals on house price dynamics. González et al. (2005) developed this model to describe heterogeneous panels, where the coefficients are allowed to vary between regions and with time.

This chapter provides empirical evidence for heterogeneity in spatial spillovers of house price developments across time and space: house price developments in neighboring regions spill over more strongly in times of increasing neighboring house prices compared to declining neighboring house prices. Moreover, heterogeneity in the effect of the fundamentals on house price dynamics cannot be detected for all variables; real per capita disposable income and the unemployment rate have a homogeneous effect across time and space.

The contribution of this chapter is to provide for the first time a joint analysis of all three spatial aspects, namely spatial dependence, spatial heterogeneity, and heterogeneity in spatial dependence. Furthermore, this chapter explicitly models heterogeneity across time and space in spatial dependence. Finally, this chapter tries to model the disposition effect using heterogeneity in spatial spillovers.

Zusammenfassung

Bei der Analyse ökonomischer Zusammenhänge auf regionaler Ebene kommt der räumlichen Dimension eine besondere Rolle zu. Dies ist der Fall, da räumliche Daten und Modelle durch zwei räumliche Effekte charakterisiert sind: Abhängigkeit und Heterogenität. Räumliche Ökonometrie ist ein Teilgebiet der Ökonometrie, bei dem räumliche Effekte explizit in ökonometrische Modelle mit einbezogen werden.

Räumliche Abhängigkeit beschreibt den Zusammenhang, dass Beobachtungen in einer Region von Beobachtungen in Nachbarregionen abhängen (Anselin, 1988). Dies steht im Kontrast zur üblichen Annahme in Querschnittsstichproben, wonach Beobachtungen aus unterschiedlichen Regionen unabhängig voneinander sind (LeSage und Pace, 2009). Räumliche Heterogenität impliziert, dass die Modellkoeffizienten in unterschiedlichen Regionen verschiedene Werte aufweisen. Dies steht wiederum im Gegensatz zur üblichen Annahme homogener Koeffizienten (Anselin, 1988).

Die Dissertation besteht aus drei Kapiteln, wobei jedes Kapitel einen anderen räumlichen Aspekt regionaler Daten analysiert. Das erste Kapitel beschäftigt sich mit räumlicher Abhängigkeit. Hierbei wird ein besonderes Augenmerk auf die Konstruktion und die jeweiligen Effekte von verschiedenen räumlichen Gewichtungsmatrizen gelegt. Das zweite Kapitel untersucht räumliche Heterogenität. Im dritten Kapitel werden räumliche Abhängigkeit, räumliche Heterogenität und Unterschiede in räumlicher Abhängigkeit in einem Modell kombiniert. Die drei Kapitel beschäftigen sich mit unterschiedlichen ökonomischen Aspekten, wobei der räumlichen Dimension jeweils eine besondere Rolle zugestanden wird.

Kapitel 1 analysiert die Bedeutung von Unternehmertumkapital für die ökonomische Entwicklung deutscher NUTS-3 (Systematik der Gebietseinheiten

für die Statistik) Regionen. Dazu wird eine neoklassische Produktionsfunktion, welche um Unternehmertumkapital erweitert wird, geschätzt (Audretsch und Keilbach, 2004). Da Unternehmertumkapital Wissensübertragungen erleichtert, den Wettbewerb erhöht und zu einer größeren Diversität an Produkten und Problemlösungen führt, sollte es die ökonomische Entwicklung beeinflussen. In diesem Kapitel wird die räumliche Dimension der Daten explizit berücksichtigt. Das bedeutet, dass nicht nur analysiert wird, inwiefern sich Unternehmertumkapital aus einer Region i auf die ökonomische Entwicklung dieser Region auswirkt, sondern auch, wie sich Unternehmertumkapital einer Nachbarregion auf die ökonomische Entwicklung der Region i auswirkt. Sowohl der Zugang zu Unternehmertumkapital als auch Wettbewerbsaspekte werden als Argumente herangezogen, warum Unternehmertumkapital einen Effekt auf die ökonomische Entwicklung von Nachbarregionen haben könnte. Es wird erklärt, warum diese Effekte sowohl positiv als auch negativ sein können.

Zur Schätzung der Effekte wird ein räumliches Durbin-Modell verwendet, da dieses die räumliche Struktur der Daten berücksichtigt. Bei der Schätzung wird der Konstruktion und der Auswahl der geeignetsten räumlichen Gewichtungsmatrix besondere Beachtung geschenkt. Die räumliche Gewichtungsmatrix bestimmt, wie stark Region i Region j beeinflusst und umgedreht. In räumlicher Ökonometrie hängen die Schätzungen und die empirischen Rückschlüsse dieser Schätzungen von der Gewichtungsmatrix ab. Die hier verwendete räumliche Schätzmethode erlaubt es nicht nur herauszufinden, wie groß der Effekt von Unternehmertumkapital einer Region auf die ökonomische Entwicklung eben dieser Region ist, sondern auch wie groß die Übersprungeffekte des Unternehmertumkapitals aus Nachbarregionen sind.

Die Schätzungen ergeben, dass Unternehmertumkapital der Region i und benachbarter Regionen einen positiven Effekt auf die ökonomische Entwicklung der Region i ausüben. Dies wird zum einen als Beleg für Wissensübersprungeffekte durch Unternehmertum sowohl in einer Region als auch zwischen Regionen gesehen. Zum anderen deuten die Ergebnisse darauf hin, dass eine Erhöhung des Wettbewerbs in Region i und j einen positiven Effekt auf

die ökonomische Entwicklung in Region i mit sich bringt. Das bedeutet, dass Nachbarregionen mit einem hohen Niveau an Unternehmertumkapital signifikant den ökonomischen Output einer Region beeinflussen. Jedoch zeigen die Schätzungen unter Verwendung verschiedener räumlicher Gewichtungsmatrizen, dass die Signifikanz des räumlichen Übersprungeffekts zu großen Teilen von eben jenen Gewichtungsmatrizen abhängt. Dies wird dahingehend interpretiert, dass sich positive und negative räumliche Übersprungeffekte von Unternehmertumkapital aufwiegen. Die Analyse bestätigt die räumliche Abhängigkeit in der Struktur der Daten. Eine Nichtbeachtung der räumlichen Abhängigkeit würde zu verzerrten Schätzern führen.

Der Beitrag dieses Kapitels besteht darin, dass zum ersten Mal die räumliche Dimension der Daten bei der Analyse des Effekts von Unternehmertumkapital auf ökonomischen Output berücksichtigt wird. Zusätzlich wird durch die Verwendung zahlreicher Gewichtungsmatrizen aufgezeigt, inwieweit die Ergebnisse von der Wahl dieser Matrix abhängen.

Dieses Kapitel basiert auf einem gemeinsamen Papier mit Konstantin A. Kholodilin.

Kapitel 2 analysiert, inwieweit der Effekt der Selbstständigenrate auf die ökonomische Entwicklung in europäischen NUTS-2 Regionen Heterogenität aufweist. Maße für Unternehmertum, wie die Selbstständigenrate treffen keine Unterscheidung zwischen sogenannten Opportunity Entrepreneuren und Necessity Entrepreneuren. Im Gegensatz zu Opportunity Entrepreneuren tragen Necessity Entrepreneure nicht zu technologischem Fortschritt bei (Acs und Varga, 2005). Da die Selbstständigenrate also keine Unterscheidung zulässt, ist es auch nicht verwunderlich, dass die meisten Studien, die den Effekt von Unternehmertum auf die ökonomische Entwicklung analysieren, zu sehr heterogenen Ergebnissen kommen.

Theoretisch kann eine hohe Selbstständigenrate ein positives unternehmerisches Umfeld oder aber auch einen Mangel an Beschäftigungsalternativen

widerspiegeln.

Wenn das positive unternehmerische Umfeld innovative Entrepreneurre anzieht, dann ist der erwartete Effekt einer hohen Selbstständigenrate auf ökonomischen Output positiv. Spiegelt eine hohe Selbstständigenrate aber einen Mangel an Beschäftigungsalternativen wider, dann ist kein oder ein negativer Effekt zu erwarten. Diese Argumentation impliziert, dass der Effekt von Selbstständigkeit auf die ökonomische Entwicklung in verschiedenen Regionen unterschiedlich ausfallen kann, je nachdem, ob das unternehmerische Umfeld innovative Entrepreneurre anzieht oder nicht.

Das Kapitel analysiert mit Hilfe einer geografisch gewichteten Regression (GWR), ob der Effekt der Selbstständigenrate auf ökonomischen Output in europäischen NUTS-2 Regionen von der Lage der Region abhängt. Die GWR wurde von Fotheringham et al. (2002) als eine Methode zum Umgang mit räumlicher Instationarität entwickelt. Die Schätzergebnisse der GWR zeigen, dass in Regionen, die einen signifikant positiven Einfluss der Selbstständigenrate auf die ökonomische Entwicklung aufweisen, die Selbstständigenrate im Durchschnitt geringer ist als in Regionen, in denen der Einfluss signifikant negativ ist.

Anschließend wird das Konzept der Gleichgewichtsrate des Unternehmertums eingeführt, um mehr über die Quelle der räumlichen Instationarität zu erfahren. Mit Hilfe dieses Konzepts wird ein Niveau der Selbstständigenrate geschätzt, von welchem an Necessity Entrepreneurre im Vergleich zu Opportunity Entrepreneurren überwiegen. Schätzungen ergeben, dass Selbstständigenraten unterhalb der Gleichgewichtsrate des Unternehmertums einen signifikant positiven Effekt auf die ökonomische Entwicklung ausüben. Selbstständigenraten oberhalb der Gleichgewichtsrate des Unternehmertums haben einen negativen Effekt auf die ökonomische Entwicklung. Dies wird als Beleg dafür interpretiert, dass Selbstständigenraten oberhalb der Gleichgewichtsrate des Unternehmertums von Necessity Entrepreneurren dominiert werden. Selbstständigenraten unterhalb der Gleichgewichtsrate des Unternehmertums implizieren ein unternehmerisches Umfeld, welches innovative Entrepreneurre anzieht.

Der Beitrag dieses Kapitels besteht darin, dass die GWR zum ersten Mal verwendet wird, um auf regionaler Ebene die heterogenen Effekte der Selbstständigengrate auf die ökonomische Entwicklung zu analysieren. Zusätzlich versucht das Kapitel, die Ursprünge der Heterogenität zu ergründen, indem es ein Niveau der Selbstständigengrate berechnet, von welchem an Necessity Entrepreneure im Vergleich zu Opportunity Entrepreneuren überwiegen.

Kapitel 3 ist eine verknüpfende Analyse der drei räumlichen Charakteristika regionaler US Hauspreisentwicklungen: räumliche Abhängigkeit, räumliche Heterogenität und Heterogenität in räumlicher Abhängigkeit. Während räumliche Abhängigkeit und räumliche Heterogenität relativ gängige Aspekte von Hauspreisentwicklungen sind, wurde der räumlichen und zeitlichen Heterogenität in räumlicher Abhängigkeit in der Literatur wenig Aufmerksamkeit geschenkt. In diesem Kapitel wird der sogenannte Dispositionseffekt herangezogen, um die Heterogenität der Übersprungeffekte von Hauspreisentwicklungen zu erklären. Der Dispositionseffekt bezeichnet das Phänomen auf den Finanzmärkten, dass Investoren Wertpapiere, die im Preis steigen, zu schnell abstoßen und jene, die im Preis fallen, zu lange halten. Bei fallenden Hauspreisen impliziert der Dispositionseffekt geringere Übersprungeffekte von Hauspreisentwicklungen zwischen Regionen als bei steigenden Hauspreisen.

In einem ersten Schritt wird in diesem Kapitel eine räumliche Panelregression geschätzt, um die globale räumliche Abhängigkeit der Hauspreisentwicklung zu analysieren. Danach wird ein sogenanntes Panel Smooth Transition Regression Model verwendet, um sowohl die Heterogenität der räumlichen Abhängigkeit als auch die Heterogenität der Effekte der erklärenden Variablen auf die Hauspreisentwicklung abzubilden. Diese Schätzmethode wurde von González et al. (2005) entwickelt, um heterogene Panels zu beschreiben und erlaubt den Koeffizienten, über die Zeit und den Raum zu variieren.

Das Kapitel liefert empirische Evidenz für die Heterogenität räumlicher Übersprungeffekte regionaler Hauspreisentwicklungen: Hauspreisentwicklungen in Nachbarregionen springen in Zeiten steigender Hauspreise stärker über als in Zeiten sinkender Hauspreise. Des Weiteren weisen nicht alle

verwendeten erklärenden Variablen heterogene Effekte auf. Die Effekte des realen Pro-Kopf-Einkommens und der Arbeitslosenquote variieren nicht über die Zeit und die Regionen.

Der Beitrag dieses Kapitels besteht in der erstmaligen zusammenhängenden Analyse der drei räumlichen Aspekte regionaler Hauspreisdaten: räumliche Abhängigkeit, räumliche Heterogenität und Heterogenität in räumlicher Abhängigkeit. Des Weiteren wird explizit die räumliche und zeitliche Heterogenität in räumlicher Abhängigkeit modelliert. Schließlich wird in diesem Kapitel erstmalig der Dispositionseffekt durch die Heterogenität räumlicher Übersprungeffekte der Hauspreisentwicklungen abgebildet.

Chapter 1

Do Regions with Entrepreneurial Neighbors Perform Better? A Spatial Econometric Approach for German Regions

*"This chapter is based on an article which is joint with Konstantin A. Kholodilin, published in REGIONAL STUDIES, published online: 16 July 2012, available online at:
<http://www.tandfonline.com/10.1080/00343404.2012.697143>."*

1.1 Introduction

Despite a vast literature about economic growth, there is no universally accepted model. The Mankiw et al. (1992) human capital augmented Solow model probably produces the most convincing empirical results. Yet, an emerging literature suggests that economic development is highly related to the abundance of small entrepreneurial firms. In this literature, start-ups constitute an important link between knowledge creation and knowledge commercialization, which will generate economic output. Acs and Armington (2004) find for the US that firm start-ups are an important vehicle, through which knowledge spills over and contributes to economic growth. Müller (2006) find the same for West German regions. The Sutter (2010) results show that the commercial introduction of knowledge via firm start-ups has a larger effect on economic

growth than pure knowledge creation. Furthermore, for West German regions in the 1990s, Audretsch and Fritsch (2002) find that start-up rates have a positive impact on growth rates. Moreover, Audretsch and Keilbach (2004) discover in their analysis of German regions that in 1992 the start-up rate had a significant effect on economic output. More generally, Fischer and Nijkamp (2009) state that regional change is the result of entrepreneurial activity where innovations play a key role.

The Knowledge Spillover Theory of Entrepreneurship (Acs et al., 2009), which underlies the aforementioned studies, focuses on individuals with endowments of new economic knowledge.¹ This knowledge is typically generated in a university or at an incumbent firm, where the individual works. The expected value of this new knowledge can be higher for this individual than for the decision maker in the university or in the incumbent firm. If the expected return is sufficiently high and the costs of starting a new business sufficiently low, the individual will enter the market and start his or her own business. The subsequently created start-up is the vehicle through which knowledge spills over from the source of knowledge production to a new firm that commercializes it.²

However, the decision to become an entrepreneur depends, as well, on a positive entrepreneurial environment. This introduces the concept of entrepreneurship capital. Audretsch and Keilbach (2004) define it "as a region's endowment with factors conducive to the creation of new business" (Audretsch and Keilbach, 2004, p. 951). These factors include, for example, individuals who are willing to start a new business, an innovative milieu, networks, institutions that help with business formation, and institutions like banks that are willing to share risks.³ In the Knowledge Spillover Theory of Entrepreneurship, the knowledge stock is only a necessary condition for economic growth. This theory gives a deeper understanding of the essential role of entrepreneurship for economic development.

However, entrepreneurship capital affects not only economic output, because

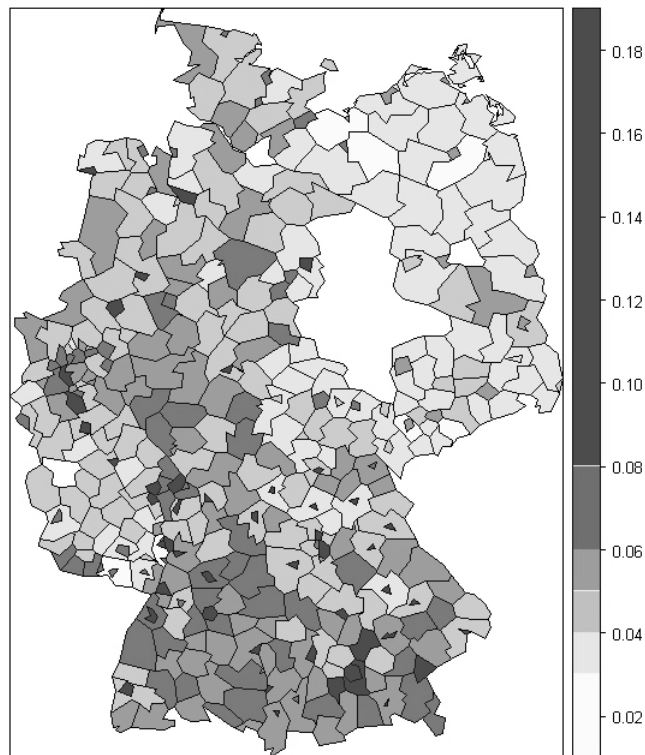
¹Economic knowledge is knowledge that holds commercial opportunity.

²This is in contrast to Lucas Jr. (1988) and Romer (1990), where knowledge exogenously spills over between firms.

³Entrepreneurship capital is what is also known in the literature as entrepreneurship culture (Armington and Acs, 2002).

it facilitates knowledge spillovers but also because it increases competition due to an increased number of entries. Fritsch and Müller (2004) argue that intensified competition increases efficiency and innovation. Furthermore, it leads to greater diversity of products and problem solutions. A greater diversity of products may be stimulating for economic development as it favors, for example, follow-up innovations. In general, these improvements in the competitiveness of an economy, together with the knowledge spillover characteristic, make entrepreneurship capital an important factor in improving economic performance.

Figure 1.1: Gross value added per working-age population, 2008



Notes: in million Euro, data available for 411 NUTS-3 regions, missing values are blank

In our analysis of regional economic growth in Germany, we take this essential role of entrepreneurship into account. Following Audretsch and

Keilbach (2004), we estimate a neoclassical production function model augmented by entrepreneurship capital. Several studies show that the effect of firm entry on regional development is distributed over a longer period of time. Audretsch and Fritsch (2002), for example, find start-up rates to be important for long-term economic development. Furthermore, Fritsch and Müller (2004) and Fritsch (2008) show that a positive effect of new business formation on economic development for German regions can be observed between five and ten years after entry and that this effect is largest eight years after entry.⁴ That is why we include the lag of entrepreneurship capital. We choose a regional context as it is widely recognized that the region is the fundamental basis of economic life (Capello and Nijkamp, 2009). In Figure 1.1 it can be seen that output in German NUTS-3 regions (Nomenclature of Territorial Units for Statistics) varies strongly across regions. Porter (2003) points out that such differences in economic performance across regions can be observed in every country. He concludes that determinants of economic performance are likely to be found as well on a regional level.

The contribution of this paper is that we take the spatial dimension of the data explicitly into account. That means that we not only analyze the effect of entrepreneurship capital in a specific region i on economic performance of that region, but we also analyze the effect of entrepreneurship capital in neighboring regions on economic performance of region i .

To estimate these spatial spillover effects of entrepreneurship capital we use a spatial Durbin model where we put special emphasis on the creation and the choice of the weight matrix. This is another aspect not receiving much attention, even in the spatial econometric literature. For the case of Germany we could not find a single paper where the choice of the weight matrix played a major role. The weight matrix determines to what extent region i affects

⁴Fritsch and Müller (2004) detect three stages in the impact of entry on economic performance. At first entry has a positive effect, but soon after the business formation some new firms fail to be competitive and some incumbent firms will leave the market as well. This negative effect of new business formation on economic development is ultimately reversed when increased competitiveness is achieved. These three stages were also found by Carree et al. (2007) for OECD countries.

region j and vice versa. In spatial econometrics, inference and estimates depend on the weight matrix used. Different weight matrix specifications will have an important impact on coefficient estimates (LeSage and Fischer, 2008). The spatial econometric estimation method allows us to not only find out how large the effect of entrepreneurship capital is on economic output in a specific region, but also how large the spillover effect is coming from neighboring regions. This will allow us to answer the question posed in the title, namely, do regions with entrepreneurial neighbors performs better?

We find entrepreneurship capital of region i and entrepreneurship capital of neighboring regions to have a positive effect on economic performance of region i . On the one hand, this is interpreted as evidence in favor of knowledge spillover effects via entrepreneurship within one region and between regions. On the other hand, we argue that improvements in the competitiveness of both region i and regions j will have a positive effect on economic output of region i . Having neighbors with a high level of entrepreneurship capital, i.e., entrepreneurial neighbors, significantly affects a region's performance. However, the significance of the spatial spillover effects largely depends on the choice of the weight matrix. We see this as evidence that positive and negative spatial spillover effects of entrepreneurship capital cancel out. The analysis confirms a spatial dependence structure. A failure to account for it would result in biased estimates.

The remainder of the paper is organized as follows. Section 1.2 explains in more detail the spatial dimension of entrepreneurship capital. In section 1.3 we present the theoretical model and the data description. Section 1.4 treats spatial estimation issues. In detail we explain the spatial Durbin model, the Bayesian estimation method, the correct coefficient interpretation and the creation and comparison of the weight matrices. Section 1.5 presents our empirical results, while section 1.6 concludes.

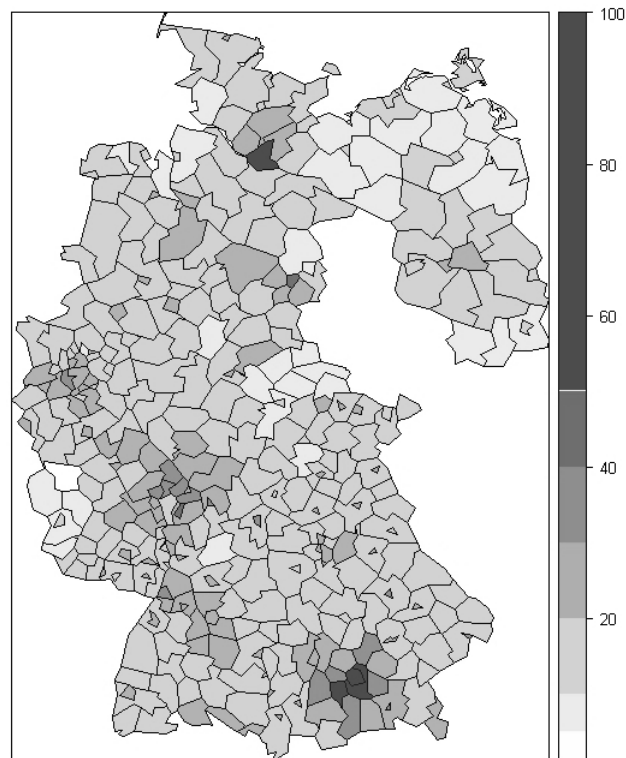
1.2 The Spatial Dimension of Entrepreneurship Capital

There are three arguments why economic performance should be analyzed taking into account the spatial dimension of entrepreneurship capital. The first two address the accessibility of entrepreneurship capital while the third covers competition issues.

For the first argument we start from the quite general theory that knowledge spillovers are more important in research and development (R&D) intense industries (Arrow, 1962). It is empirically shown that the knowledge in those innovative industries exhibits a certain degree of tacitness (Doering and Schnellbach, 2006). The difficulty in communicating such knowledge makes direct interaction necessary. That implies that space matters for knowledge to spill over because the cost of face-to-face contacts increases with distance. This is supported by the finding of Audretsch and Feldman (1996) that innovative activity tends to cluster spatially. For the Knowledge Spillover Theory of Entrepreneurship this suggests that the entrepreneur's decision on where to set up the new business depends on the location of the university or firm that he or she comes from. Locating near the knowledge source will allow the maintenance of their personal social network. These face-to-face contacts are an opportunity to exchange information on research advances, profitability, and entrepreneurial opportunities (Doering and Schnellbach, 2006). Furthermore, Audretsch and Lehmann (2005) argue that especially in the case where the knowledge source is a university it makes sense to locate in close proximity as this facilitates access to a large pool of skilled labor. Empirical evidence supports this theory. Audretsch and Lehmann (2005) find that start-ups in Germany tend to cluster within close proximity to the knowledge source. Similarly, Anselin et al. (1997) find that for US metropolitan statistical areas, university knowledge spillovers extend over a range of only 50 miles (about 80 km) from the area of innovation. Funke and Niebuhr (2005) find a range of 120 km for German NUTS-3 regions. Given the fact that on average the German NUTS-3 regions have a reach of about 32 km, setting up a new business in proximity to the knowledge source also

includes neighboring regions. It is the accessibility of entrepreneurship capital (in this case especially the network) that is decisive. Therefore, choosing a specific location depends not only on the entrepreneurship capital of that region but also on entrepreneurship capital of the neighboring region where the knowledge source is located.

Figure 1.2: Start-up rates in knowledge intense areas, 1997-2004



Notes: per 10,000 working-age persons, data available for 385 NUTS-3 regions, missing values are blank

The second argument starts from the observation that regions with large amounts of innovative activity experience a high level of knowledge-based entrepreneurial activity and that innovative activity tends to cluster spatially (Audretsch and Feldman, 1996). This in turn means that knowledge-based entrepreneurial activity clusters as well. This clustering is *per se* conducive

to the creation of new businesses.⁵ Figure 1.2 demonstrates that knowledge intense start-ups cluster, for example, around Munich and in the Rhine-Main area. Porter (1998a) describes that an individual working within a cluster can more easily discover entrepreneurial opportunities. Furthermore, a cluster facilitates access to employees, suppliers, and specialized information. The author further argues that the local financial institutions are already familiar with the cluster. This implies that they have a better understanding of the risk involved and therefore offer better conditions. These factors will make it more likely that a new business will locate within a cluster. This is proven empirically. Both Rocha and Sternberg (2005) for Germany and Delgado et al. (2010) for the US find a positive impact of clusters on entrepreneurship. Given the definition of entrepreneurship capital, the characteristics of a cluster can very well be summarized under this notion. Regarding the cluster size, Porter (1998a) argues that clusters may cross state or even national boundaries. Our unit of analysis is much smaller than that. The decision on where to locate is not only based on the characteristics, i.e., entrepreneurship capital of the region i but also on characteristics of the cluster and clusters can be spread across several regions. Again, the accessibility of entrepreneurship capital is decisive, and accessibility is not limited to the region of location.

The third argument for a spatial dimension of entrepreneurship capital is that it increases competition because of a higher number of entries. Of course, competitors can not only be found in the NUTS-3 region i , where the new entrepreneur is located in but also in neighboring or even geographically distant regions. Increased competition will lead to improvements in the competitiveness and has therefore a positive effect on economic development of region i . This explains another positive effect of entrepreneurship capital of the neighboring regions on economic performance of region i . But these improvements in the competitiveness can also have negative effects. If the firms of a specific region are especially competitive because of a high level of entrepreneurship

⁵Porter (1998b) defines a cluster as "a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities" (Porter, 1998b, p. 199).

capital, entry barriers in geographically distant regions, where there is no access to this entrepreneurship capital are also high. Potential entrepreneurs will not decide to start their own business and some firms may even leave the market due to the lack of competitive advantages. In this case entrepreneurship capital of other regions, $j \neq i$, could negatively affect the economic performance of region i .

To sum up, we provide theoretical foundations for both positive and negative spatial spillover effects of entrepreneurship capital.

1.3 Model

To analyze the knowledge spillover effects via entrepreneurship, we consider a neoclassical production function, which not only includes the standard variables physical capital (K) and human capital (H) as explanatory variables but also entrepreneurship capital (E): Equation (1.1). In this way we follow the approach of Audretsch and Keilbach (2004).

$$Y = F(K, H, E, L) \tag{1.1}$$

The variables are divided by labor (L) so we work with variables expressed per effective unit of labor and have thereby productivity expressions, y , k , h , e . With a Cobb-Douglas specification of the production function we get:

$$y_i = a k_i^{\alpha_K} h_i^{\alpha_H} e_i^{\alpha_E}, \tag{1.2}$$

where $i = 1, \dots, n$ denotes regions and a represents the state of the technology. Taking logs yields:

$$\ln y_i = \ln a + \alpha_K \ln k_i + \alpha_H \ln h_i + \alpha_E \ln e_i. \tag{1.3}$$

We estimate this theoretical model using data for 337 German NUTS-3 regions for the year 2008.⁶ All variables are ratios per working population.

The dependent variable economic output, y , is measured by gross value added at basic prices; physical capital, k , is calculated with the perpetual inventory method. This procedure allows computation of the stock of physical capital (K) as the weighted sum of past investments (I) in manufacturing and mining ($K_t = I_t + (1 - \delta)K_{t-1}$). For the calculation, we chose data on investment in 1995 as initial capital stock. We assume a depreciation rate, δ , of five percent (Barro and Sala-i-Martin, 1995; Chew and Tan, 1999). We know that this captures only part of the total investments and could result in misleading coefficients, but data on gross fixed capital formation are not published at the regional level. Further, we use the share of employees with technical college or university degree in the working age population to measure human capital, h . This definition is in line with what is used in the literature (Barro and Lee, 1993; LeSage and Fischer, 2008; Fischer et al., 2009). Following Audretsch and Keilbach (2004), entrepreneurship capital, e , is approximated by start-up rates in knowledge intensive areas. We use this variable since entrepreneurship capital is a variable that cannot be observed but should manifest itself through high start-up rates. The *Mannheim Enterprise Panel* (MUP) provides those start-up rates for several categories. We defined six categories as being knowledge intensive and aggregated them (Table 1.4). We only use knowledge intensive start-ups as we are interested in knowledge spillovers via entrepreneurship, which are most likely to occur in these areas (Acs et al., 2009). As there are years where the number of new business formation in some knowledge intense areas is very small, the data are published for a time span of four years. In particular, data are available for 1997-2000, 2001-2004, and 2005-2008. As mentioned above, a positive effect of new business formation on economic development can be observed between five to ten years after entry. That is why we aggregated the start-up rates for 1997-2000 and 2001-2004. This should allow us to best capture the positive effects of entrepreneurship capital.

⁶We limit our analysis to 337 out of 413 NUTS-3 regions (in 2008) due to limited data availability. In particular, the data on investments are not complete. The 337 regions are plotted in Figure 1.3, a complete list is provided by the authors upon request.

1.4 Spatial Econometric Modeling

1.4.1 Spatial Durbin Model

As LeSage and Pace (2010) note, data collected from regions are often not independent. This spatial dependence requires a special estimation method because neglecting this structure would result in biased estimates. There are four well known spatial econometric specifications, namely the spatial lag model (SAR), which includes a spatial lag of the dependent variable, the spatial error model (SEM), which includes a spatial lag in the error term, the spatial autoregressive moving average model (SARMA), which includes a spatial lag of the dependent variable and in the error term, and the spatial Durbin model (SDM), which includes a spatial lag of the dependent and the explanatory variables.

The spatial lag $\sum_{j=1}^n W_{ij}y_j$ is the weighted average of the spatially lagged variables of the neighboring regions. W is the spatial weight matrix of dimension $n \times n$, where n is the number of regions. If two regions i and j are spatially related, the element $w_{ij} \neq 0$, otherwise $w_{ij} = 0$. By convention, a region cannot be a neighbor to itself, $w_{ii} = 0$. To simplify interpretation the weight matrix is usually row-standardized, so that the row sums are equal to one. The spatial lag operator then corresponds to the weighted average of neighboring observations.

LeSage and Pace (2009) point out that the SDM is the only model that will produce unbiased estimates no matter which of the mentioned data generating processes is underlying. This is why we choose the spatial Durbin model as appropriate estimation specification (Equation (1.4)). This model further nests the spatial lag and the spatial error model, i.e., models involving dependence in the error term and in the dependent variable.

$$\begin{aligned} y &= \alpha_n + \rho W y + X\beta + WX\gamma + \epsilon \\ \epsilon &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{1.4}$$

In this equation, ρ measures the strength of the spatial lag dependence of the dependent variable, Wy . γ measures the strength of the spatial lag dependence of the explanatory variables, WX . This spatial model specification, applied to our neoclassical production function model, yields the following equation that we estimate.

$$\begin{aligned} y &= \alpha_n + \rho Wy + \beta_1 k + \beta_2 h + \beta_3 e + Wk\gamma_1 + Wh\gamma_2 + We\gamma_3 + \epsilon \quad (1.5) \\ \epsilon &\sim N(0, \sigma^2 I_n) \end{aligned}$$

Where y is the dependent variable economic output, Wy is the spatial lag of economic output, Wk is the spatial lag of the independent variable physical capital, Wh is the spatial lag of the independent variable human capital, and We is the spatial lag of the independent variable entrepreneurship capital. This specification allows us not only to explicitly account for spatial dependence in the data but also to gain insight into the regional spillovers of the three explanatory variables.

1.4.2 Estimation Method

Spatial models can be estimated using maximum likelihood or Bayesian estimation methods. Our focus is on the comparison of different weight matrices. Tests, like the likelihood ratio test, that use the log-likelihood function values to compare the models can only be used with nested models. If two models have different weight matrices, then they cannot be considered as nested. This is why we use Bayesian estimation. The Bayesian posterior model probabilities allow model comparison even for non-nested models.⁷

Bayesian estimation in general is centered on posterior probabilities. $P(\theta|D)$ is the so called posterior probability of the parameters, $\theta=(\alpha, \beta_r, \gamma_r, \rho, \text{ and } \sigma^2)$, given the data, D , and reflects the belief about the parameters after collecting

⁷We use Matlab codes by LeSage to estimate our spatial models.

the data.

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (1.6)$$

The posterior distribution represents an update of the prior distribution given the data. $P(D|\theta)$ is the model likelihood and $P(\theta)$ is the prior distribution of the parameters that reflects previous knowledge or uncertainty prior to observing the data. The probability of the data $P(D)$ is not of great interest as it does not involve the parameters $\theta=(\alpha, \beta_r, \gamma_r, \rho, \text{ and } \sigma^2)$. Bayesian inference about parameters is entirely based on the posterior distribution $P(\theta|D)$.

We apply the Bayesian Markov Chain Monte Carlo approach to estimate the parameters $\alpha, \beta_r, \gamma_r, \rho, \text{ and } \sigma^2$, where r is 1 to 3, and stands for the explanatory variables. By applying this method, we work with a large random sample from the posterior distribution and not with the precise analytical form of the density. A large sample of the posterior probability distribution allows us to approximate the analytical form of the probability density. We follow LeSage and Pace (2009) and assign the normal prior to α, β , and γ , the inverse gamma prior to σ^2 , and the uniform prior to ρ .

We sample 7500 times from the conditional distribution and assume that the sampler achieves its steady state after 2500 draws. The last 5000 draws are interpreted as coming from the posterior distribution. We use the large sample of parameter draws from the posterior distribution to make inference about $\alpha, \beta_r, \gamma_r, \rho, \text{ and } \sigma^2$.

We also account for heteroscedasticity in the data by extending the above described Markov Chain Monte Carlo estimation by variance scalars that can accommodate non-constant variance of the error term:

$$\epsilon \sim N(0, \sigma^2 V), \quad (1.7)$$

where V is a diagonal matrix containing the parameters (v_1, v_2, \dots, v_n) , which are unknown and need to be estimated. We assign a chi-squared prior distribution, $\chi^2(s)/s$, to the v_i terms. The Markov Chain Monte Carlo sampling scheme is extended by an additional conditional distribution for the variance scalars. Following LeSage and Pace (2009), s of the chi-squared

prior distribution is set to four, as this is consistent with a prior belief in non-constant variance and outliers.

1.4.3 Coefficient Interpretation

The coefficients β and γ of Bayesian estimation cannot be interpreted as marginal effects. This comes from the spatial dependence structure in the data. A change in the explanatory variable of region i will affect the region i itself, which is called a direct impact, and potentially also the neighboring regions, which is called an indirect impact. Since the neighboring regions affect region i as well, there is feedback in the system. Spatial econometric models are able to capture this effect. LeSage and Pace (2009) explain how to calculate these summary marginal measures of impact. Using their method we are able to calculate the direct, the indirect, and the total average impact effects of the variables. The average direct effect is the one that comes from the same region i . The average indirect effect or spatial spillover effect is the one that comes from the other regions $j \neq i$. The average total effect is the sum of the direct and the indirect effect.

During the Markov Chain Monte Carlo sampling we can construct these three summary measures. We simply use at each pass through the Markov Chain Monte Carlo sampling loop the sampled parameters α , β , γ , and ρ to calculate the direct effect, the total effect, and by subtracting the direct effect from the total effect, the indirect effect.⁸ Then we can construct the entire posterior distribution for the three types of marginal effects using the 5000 saved draws.

1.4.4 Spatial Weight Matrix Comparison

We choose the spatial Durbin model, but we still cannot simply start our estimation as we do not know which weight matrix we should use. There are

⁸Technical details can be found in the working paper version of this paper.

several weight matrix specification possibilities. The weights could be based on geographical, technological, economic, demographical, or political distance. Finding the weight matrix that best reflects the spatial dependence is a key element of spatial econometric analysis. As previously mentioned the accessibility of entrepreneurship capital plays an important role for spatial spillover effects of entrepreneurship capital. This accessibility is limited by distance. In general, there is a common agreement that knowledge spillovers are favored by face-to-face contacts and interpersonal relationships. Therefore, they are as well limited by distance (Funke and Niebuhr, 2005; Baldwin et al., 2008). We conclude that the use of geographical distance in our analysis is appropriate. But ties between regions are not only determined by geographical distance. Economic interaction of regions depends as well on non-geographic distance. Autant-Bernard et al. (2007) analyze knowledge transfers through R&D and find that social network effects matter for the collaboration between firms. Furthermore, Scherngell and Barber (2009) find that firms collaborate more if they are close in technological space, and it is not necessarily dependent upon geographical space. Those results imply that knowledge spillovers do not have to be regionally bounded. We argued earlier that entrepreneurship capital improves competitiveness through greater competition, where the competitor could also locate in a geographically distant region. In this case entrepreneurship capital of the region of a technological close competitor has an effect of economic performance of the unit of observation. To capture these spatial spillovers we need to take non-geographical measures of distance into account.

The weight matrix based on geographical proximity can be constructed using either of two approaches: a binary measure of contiguity or a continuous measure of distance. Contiguity measures take the word neighbor in its narrowest sense. If two regions i and j share a common border they are considered to be first-order contiguous and the value of 1 is assigned to w_{ij} . Higher orders of contiguity can be considered as well. Contiguity of order c assigns the value of 1 to regions which share a common border with a region that is a $(c - 1)$ order contiguous region.

Continuous distance measures assign a value $w_{ij} > 0$ to regions that are within a certain distance from region i . This value is obtained using, for example, geographical coordinates of two regions i and j together with a distance decay function. Distance decay functions have the effect of reducing the influence of regions j on region i as the distance between them increases. When using distance measures most studies rely on geographical proximity. LeSage and Fischer (2008) use geodesic distance, road travel time distances for cars, and drive time distances for heavy goods vehicles. Parent and LeSage (2008) do not only rely on geographical distance to create their weight matrix, but also on technological distance. They measure the distance between technological fields by using patent activity occurring between regions in the same field of technology defined by the International Patent Classification. LeSage and Polasek (2008) incorporate prior information about commodity flows transported by road and rail into the spatial connectivity structure, whereas Beck et al. (2006) use volume of trade flows. For the United States, Case et al. (1993) construct weight matrices based on economic and demographical proximity, where they use per-capita income for economic proximity and the percentage of the population that is black for demographical proximity. Bhattacharjee and Jensen-Butler (2006) propose to use the data in the model to estimate the spatial weight matrix that is consistent with the observed pattern of spatial dependence.

For our analysis, we create 56 row standardized spatial weight matrices that can be arranged by four types of geographical and two types of non-geographical distance measures.

For the first nine weight matrices we use the c nearest neighbors, where $c = 1, \dots, 9$. If a district is a c nearest neighbor, a value of one is assigned to that region. This procedure results in a binary contiguity matrix.

The direct distance, calculated between the centers of two regions together with the cut-off distance b is used for the next 14 weight matrices. The cut-off distance $b = 30$ determines that only regions that are within a 30 km radius around the region under examination have an impact. We choose b to be 30, 50, 100, 150, 200, 250, and 300 and created a weight matrix for each distance

twice: once with the power distance decay function ($w_{ij} = \frac{1}{d_{ij}^p$, where we set p equal to one, and d_{ij} is the distance between region i and region j) and once with the exponential distance decay function ($w_{ij} = \frac{1}{\exp^{pd_{ij}}}$, again, we set p equal to one).

The next 14 matrices are created similarly to the previous 14, but using the road distance between regions instead of the direct distance. We choose the most fuel-efficient route, not the shortest or the fastest one. Even though the infrastructure is quite well developed in Germany, trucks often need to cover longer distances on the road than the direct distance. For instance, the direct distance between Flensburg and Lübeck, which are both cities located in northern Germany, is 132 km. The road distance, taken from the route planning site viamichelin.de, is 160 km. If we take one city in the north of Germany and another in the south the difference between direct and road distance becomes more pronounced in absolute terms. The direct distance between the aforementioned Flensburg and Lindau, a city on the Lake of Constance, is 800 km, whereas the road distance is 948 km.

The following twelve weight matrices account for the time on the road that one needs to cover the distance between two regions. As cut-off time d , we use $d = 1, \dots, 6$ hours. The duration may give us additional insight as it varies even if we have quite similar distances. The road distance from Flensburg to Göttingen is 419 km, while that from Flensburg to Grafschaft Bentheim is 412 km. In the first case, the duration of trip is 5 hours and 30 minutes, for the second it is only 4 hours and 15 minutes.

For the weight matrices representing technological distance we did not follow the Parent and LeSage (2008) approach.⁹ Instead, we use the number of employees attributed to 18 different branches of economic activity to create a measure of how close the regions are in terms of technology. If regions have a similar employment structure we assume that they have similar technologies. In that case, even if the geographical distance is large, spillovers are more likely and accordingly a larger weight is assigned. In detail, we constructed a

⁹Unfortunately, we do not have access to this data.

Euclidean distance:

$$w_{ij} = \sqrt{\left(\sum_{k=1}^K (F_{ki} - F_{kj})^2\right)}, \quad (1.8)$$

where F_{ki} is the number of employees in the area of economic activity k ($k = 1, \dots, K$, with $K = 18$) in region i . This measure of distance varies between 0.47 and 14.03. The two technologically closest regions would get a weight equal to $\frac{1}{0.47}$. Again we use different cut-off distances, namely 2, 3, 4, 6, 9, 12.¹⁰

For the last weight matrix we use data on commuting in- and out-flows.¹¹ Commuting data are especially interesting because they give evidence about how open and mobile regions are (Patuelli et al., 2009). Furthermore, they allow analyzing the accessibility of a region (Reggiani et al., 2011). This is especially interesting for our analysis, since we argue in the introduction that the accessibility of entrepreneurship capital is decisive. Only if a region is easily accessible there is an opportunity for spatial interaction. In general, the interaction between regions represented by commuting flows is likely to result in spatial dependence (Badinger and Url, 2002). So it is a straightforward way to model the regional spillovers using these data. For the weight matrix we have the place of work in the rows and the place of residence in the columns of the matrix. This construction implies that the more people of a certain residence region, j , are working in region i , the greater are the spillovers from the area of residence to the area of work. We standardized the absolute numbers through dividing them by the row sums. Therefore, the typical element of this asymmetric matrix is the share of commuters coming from region j to i in the total number of commuters coming to region i . The average commuting distance is below 20 km. That implies that this matrix captures only very close spillovers.

It should be mentioned that the technological weight matrices and the matrix created using commuting flows are probably not exogenous to the model.

¹⁰We also tried the measure proposed by Jaffe (1986) $\left(\frac{\sum_{k=1}^m F_{ki}F_{kj}}{(\sum_{k=1}^m F_{ki}^2 \sum_{k=1}^m F_{kj}^2)^{1/2}}\right)$. The results are close to our measure.

¹¹We thank Franz-Josef Bade for kindly providing these data.

Commuting flows, for example, not only represent the economic ties between two regions but they are likely to be determined by the economic activity of a region, y_i . When the weights are not exogenous, the model becomes non-linear with endogeneity (Anselin, 2002). We are aware of this problem but we try these matrices anyway as we want to deliver a complete picture on how the results depend on the different geographical and non-geographical weight matrices.

Bayesian model comparison allows us to determine which of the 56 weight matrices best fits the data. Therefore, we look at the posterior model probabilities. But first we need to specify prior probabilities for each model. We assign to each model the same probability, namely $\frac{1}{m}$, where m is the number of different models, namely 56. Together with the prior distributions for the parameters we can calculate posterior model probabilities. Those posterior probabilities are then directly compared. The model with the highest posterior probability fits the sample data best.

1.5 Estimation Results

1.5.1 Coefficient Comparison

The theoretical discussion above points out that the choice of the the appropriate weight matrix is an important part in the estimation of spatial models. The marginal effects will depend on the choice of this matrix. That is why in a first step we estimate the spatial Durbin model for each of the 56 different weight matrices and compare the marginal effects. Table 1.1 shows that the direct, the indirect, and the total marginal effect estimates of the spatial Durbin models differ depending on the choice of the weight matrix. This is a result also found by Harris et al. (2011) for the UK. We present here the minimum and the maximum value of those coefficients that were significant in

the estimation at least at a ten percent level.

Table 1.1: Coefficient comparison

Dependent variable: economic output, y			
	Direct Effects		
	Min	Max	Significant Coefficients
physical capital k	0.10	0.18	56
human capital h	0.22	0.34	56
entrepreneurship capital e	0.06	0.25	54
	Indirect Effects		
	Min	Max	Significant Coefficients
physical capital k	0.08	0.53	33
human capital h	-2.08	-0.12	38
entrepreneurship capital e	0.06	0.63	15
	Total Effects		
	Min	Max	Significant Coefficients
physical capital k	0.22	0.63	36
human capital h	-1.1	0.14	28
entrepreneurship capital e	0.23	0.77	36

Note: The spatial Durbin model was estimated for each of the 56 different weight matrices. The minimum and the maximum value of those coefficients that were significant in the estimation at least at a ten percent level are presented.

The coefficients of the direct marginal effects do not vary much across models. This can also be seen in the figures 1.4, 1.5, and 1.6, where we plot the direct, the indirect, and the total effects of physical capital (k), human capital (h), and entrepreneurship capital (e) for all estimations. The dot stands for the total marginal effect. Again, only coefficients that are significant are presented.

The size of the spatial spillover effect varies quite strongly for the three variables over the different estimations. This can be seen from the smallest and the largest coefficients in Table 1.1 and from Figures 1.4 through 1.6. For

physical capital, we find the largest spatial spillover effects in those models where the weight matrices are constructed based on the direct or road distance with a cut-off point of 100 and 150 km. For human capital, large negative spatial spillover effects are found for road and direct distances up to 200 km or two hours driving distance. It appears that regional spillovers of physical capital are especially pronounced if close regions have a weight unequal to zero in the weight matrix. Spatial human capital spillovers are strong even over little longer distances.

For entrepreneurship capital, the largest spatial spillover effect is found for the weight matrix using the closest technological distance. Furthermore, the direct distance with a cut-off value of 100 km and the commuting matrix deliver large spatial spillover effects. This result suggests that both technological and geographical proximity results in pronounced spatial spillover effects of entrepreneurship capital. Furthermore, the importance of accessibility of entrepreneurship capital is underlined, since the commuting matrix, which mirrors how accessible a region is, delivers large spatial spillover effects. The results further suggest that finding a positive effect of entrepreneurship capital across regions depends to a large extent on the choice of the weight matrix. Only 15 out of 56 coefficients of the indirect effects of entrepreneurship capital are significant.¹²

1.5.2 Model Comparison

These different coefficient results emphasize the importance of carefully choosing the weight matrix. We elaborated above, that if the weight matrix does not represent the true spatial dependence structure we could possibly get misleading results. As described, we use measures of physical and technological distance to create the weight matrices. As described in section 1.4.4, posterior model probabilities are then compared to get the most appropriate model specifications. Those can be found in Table 1.2, where the seven highest prob-

¹²Boxplots used to visually summarize the distribution of coefficients of the different estimations can be found in the working paper version of this paper.

abilities are reported.

Table 1.2: Posterior model probabilities

Distance	Proba
Direct distance, cut-off 50 km, power distance function	0.646
Direct distance, cut-off 100 km, power distance function	0.277
Duration distance, cut-off 2 hours, exponential distance function	0.027
Road distance, cut-off 100 km, power distance function	0.024
Duration distance, cut-off 2 hours, power distance function	0.021
Road distance, cut-off 150 km, power distance function	0.002
Commuters	0.001

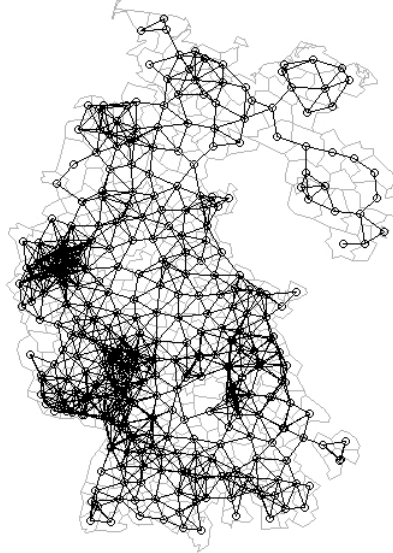
Note: Only the seven highest model probabilities are presented.

The Bayesian model probabilities point with 0.65 to the direct power distance with a cut-off distance at 50 km. Therefore, this weight matrix best fits the sample data. The links between regions implied by the chosen weight matrix can be seen in Figure 1.3. The spatial Durbin model estimation, on which we rely, is the one that uses the weight matrix based on direct power distance with a cut-off distance at 50 km. We further see that the three models with the highest probability have a spatial dependence structure where only quite close regions, namely within 100 km or two hours driving distance, exert a spatial effect.

1.5.3 Results of the Final Model

The results of the final model can be found in Table 1.3. For the sake of comparability we also present the simple OLS results in this table. We see that the Bayesian coefficient estimates are quite similar to the direct summary effects. Moreover, we see that we cannot interpret the coefficients of spatially lagged explanatory variables (Wk , Wh , We) as spillovers. The true regional

Figure 1.3: The links between regions implied by the best weight matrix



Notes: (direct distance with cut-off distance at 50 km), plotted for the 337 NUTS-3 regions used in the estimation, missing values are blank

spillovers are represented by the indirect effects and are quite different. Since the spatial spillover effect is significant for all three variables, it implies the necessity of accounting for spatial effects. Furthermore, ρ , which describes the strength of the spatial dependence of the dependent variable, is large and significant. If we would neglect the spatial dependence structure, which is clearly present, we would get biased estimates.

For physical capital all three impact measures are positive and significant. The spatial spillover effect is more than twice the magnitude of the direct effect. This is a result often found in the literature (LeSage and Pace, 2009; Parent and LeSage, 2010) and it comes from the fact that the indirect effect is the result of the accumulation of the indirect effects of several neighboring regions. LeSage and Pace (2009) show that the indirect effects from individual neighbors are smaller than the direct effect, but that the accumulation of these individual indirect effects ends up in a larger impact than the direct effect. LeSage and Fischer (2008) find a small negative

Table 1.3: Estimation output OLS and of the final Bayesian spatial Durbin model

Dependent variable: economic output, y					
Variable	Coefficients		Effects		
	OLS	SDM	Direct	Indirect	Total
constant	-3.32***	-2.42***			
physical capital $k(\beta_1)$	0.18***	0.10***	0.12***	0.27***	0.39***
human capital $h(\beta_2)$	0.27***	0.35***	0.31***	-0.49***	-0.18***
entrepreneurship capital $e(\beta_3)$	0.23***	0.09***	0.10***	0.21**	0.31***
$Wk(\gamma_1)$		0.07***			
$Wh(\gamma_2)$		-0.43***			
$We(\gamma_3)$		0.05			
ρ		0.56***			

Notes: *** Statistically significant at a one percent level, ** Statistically significant at a five percent level, W = direct distance, cutt-off distance 50 km, power distance function

indirect effect of physical capital but do not give further explanations for their result. As all variables are in logs, the total effect of 0.39 implies that a 10 percent increase in all regions physical capital stock will increase output by 3.9 percent. If we would not consider the effect of the neighboring regions in our estimation, we would largely underestimate the marginal effect of physical capital. The marginal effect of the OLS estimation is only 0.18.

The largest direct effect is found for human capital. A ten percent increase in human capital of region i will increase economic output of region i by three percent. LeSage and Fischer (2008) and Fischer et al. (2009) find a smaller direct effect of human capital for European regions, with the magnitude of 0.12, in their analysis of regional labor productivity and regional growth, respectively.

To the contrary, the indirect marginal effect of human capital is negative. That means that a ten percent increase in human capital of the regions, which according to the weight matrix have an influence on region i , will

reduce economic output of region i by almost five percent. Compared to the LeSage and Fischer (2008) and Fischer et al. (2009) studies this effect is quite large. The authors find an indirect effect of -0.11 and -0.20, respectively. The difference in the amplitude of the indirect effect is probably due to the use of NUTS-2 regions in their studies and NUTS-3 regions in ours. This implies that the indirect effect of human capital is more pronounced if smaller regions are analyzed. Furthermore, it is possible that human capital spillovers between regions are larger in a national context than in an international one. The literature mentions two explanations for a negative spillover effect of human capital. According to Fischer et al. (2009) the relative regional advantages in human capital matter the most for labor productivity. This means that an increase in neighboring human capital will result in a worse relative position of the region under consideration. Olejnik (2008) states that migration of educated workers between neighboring regions is the main reason for a change in the level of human capital. This implies that an increase in the level of human capital in neighboring regions comes from a decrease in the level of human capital in the region under consideration and has therefore a negative impact on economic performance of that region. Of course, working conditions are a crucial factor for migration decisions. Better conditions attract educated workers from neighboring regions and will consequently result in an output reduction of those regions (Nistor, 2009). Due to this large negative indirect effect, the total effect for human capital is negative as well. This result is not in line with the studies of LeSage and Fischer (2008) and Fischer et al. (2009), who find insignificant total effects of human capital. Again, if we neglect the spatial spillover effects we get misleading results. The marginal effect of human capital in the OLS estimation is 0.27. Our results suggest that regions should have a large interest in improving working conditions as there is high competition over human capital. This conclusion would not be drawn from the OLS results.

The coefficient of the direct marginal effect of entrepreneurship capital (0.10) points to the presence of knowledge spillovers via entrepreneurship as well as improved competitiveness through entrepreneurship capital. We

argued in the introduction that these two factors make entrepreneurship capital an important factor in promoting economic development. A ten percent increase in entrepreneurship capital will result in an increase in output by one percent. This is in line with the Audretsch and Keilbach (2004) results. They find a coefficient of entrepreneurship capital in a neoclassical production function model of 0.12, but they do not account for spatial dependence in their analysis of West German regions. To our knowledge, the only study, that uses entrepreneurship in a spatial Durbin model is the one by Sutter (2010). He finds a direct effect of entrepreneurship on factor productivity of 0.5 for US states.

We note in the introduction that there is theoretical evidence for a spatial dimension of entrepreneurship capital. Our final estimation delivers empirical evidence for this spatial dimension as the indirect effect of entrepreneurship capital is significant. Compared to the direct effect, the spatial spillover effect (0.21) is quite large, but it is smaller than the indirect effect of entrepreneurship (0.66) found by Sutter (2010) for the US. We conclude, on the one hand, that the accessibility of entrepreneurship capital leads the decision on where to locate. As we argued earlier, accessibility is not limited to the region of location but it also includes neighboring regions. This should be especially true for the NUTS-3 level, where the average extension is only 30 km. In this sense, knowledge spills over not just from firms or universities to the start-up but also between regions. On the other hand, this result is again evidence for competitiveness improving effects of entrepreneurship capital. Competitors not only locate in the same region as the new start-up but also in neighboring regions. The greater the entrepreneurship capital in neighboring regions, the stronger the competitors in those regions are, and the greater are the competitiveness improving effects. The more competitive a start-up is, the larger the increase in economic output in the region of location.

To answer our initial question we can say that regions with entrepreneurial neighbors perform better. However, this result largely depends on the weight matrix. In Figure 1.6 it is shown that only 15 out of 56 spatial spillover effects of entrepreneurship capital are significant. We argue above that the competitiveness improving effects of entrepreneurship capital are an entry barrier for

firms in regions with no access to this entrepreneurship capital. This negative effect offsets the positive effects mentioned above and therefore, in most estimations the spatial spillover effect of entrepreneurship capital is not significant.

1.6 Conclusion

In our analysis of regional economic performance of German regions, we estimate a neoclassical production function with physical capital, human capital, and entrepreneurship capital as explanatory variables. We are especially interested in the role of firm start-ups as link between knowledge creation and knowledge commercialization and in the competitiveness improving effect of entrepreneurship. We use the spatial Durbin model in our estimation because it allows us to take into account the spatial dependence structure of the data. We put special emphasis on the creation and comparison of diverse weight matrices. A weight matrix determines to what extent region i affects region j and vice versa. The weight matrix, which was created based on the direct distance using a cut-off distance of 50 km and the power distance decay function, is found to be the matrix that best mirrors the true spatial dependence structure. We find this result by comparing posterior model probabilities, which we calculate using a Bayesian Markov Chain Monte Carlo estimation.

We find positive direct and spatial spillover effects of entrepreneurship capital. Given our best spatial weight matrix, spatial spillovers extend over a range of 50 km from the focus region. We conclude that the pure creation of knowledge is a necessary condition for higher economic output but in order to commercialize this knowledge entrepreneurship capital is needed. Furthermore, knowledge spills over from the knowledge source to the start-up not only within a region but also across NUTS-3 borders. This is the case as the accessibility of entrepreneurship capital is decisive. In addition, the competitiveness improving effects are not limited to a certain region but are also spread across NUTS-3 borders. Therefore, entrepreneurship capital of neighboring regions and in a narrower sense entrepreneurial neighbors matter for economic performance

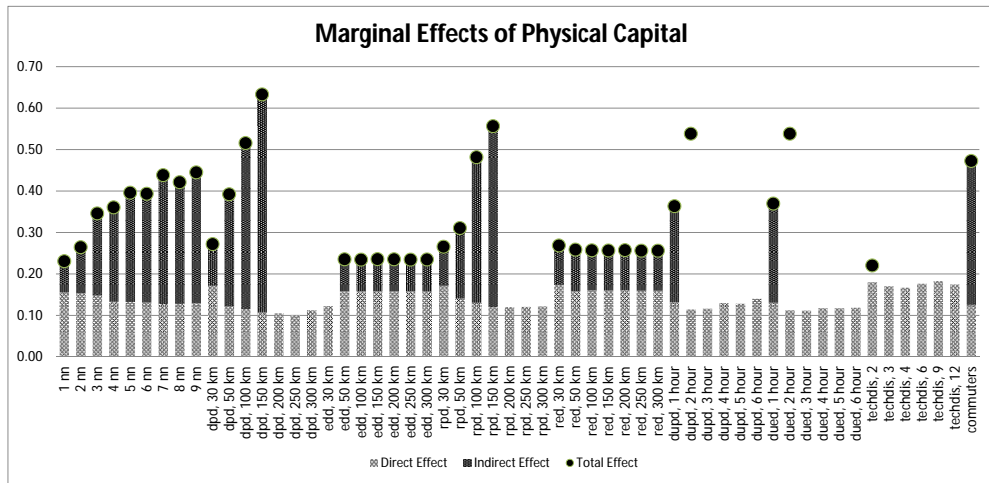
of a certain region. However, the significance of the spatial spillover effects largely depend on the choice of the weight matrix. We see this as evidence that positive and negative spatial spillover effects of entrepreneurship capital cancel out.

This analysis provides evidence for the importance of entrepreneurship capital for economic performance. Regarding the econometric setting we have shown the necessity of the use of a spatial econometric model and of a careful choice of the weight matrix.

1.A Appendix

1.A.1 Figures

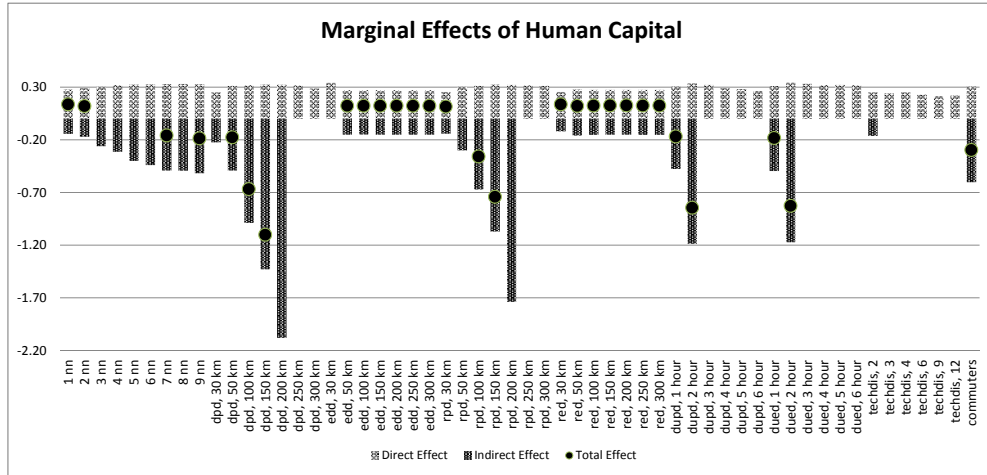
Figure 1.4: Coefficient comparison for different weight matrices: physical capital



Notes: Coefficients are only plotted if significant, nn = nearest neighbor, dpd = direct power distance, edd = exponential direct distance, rpd = road power distance, red = road exponential distance, dupd = duration power distance, dued = duration exponential distance, techdis = technological distance

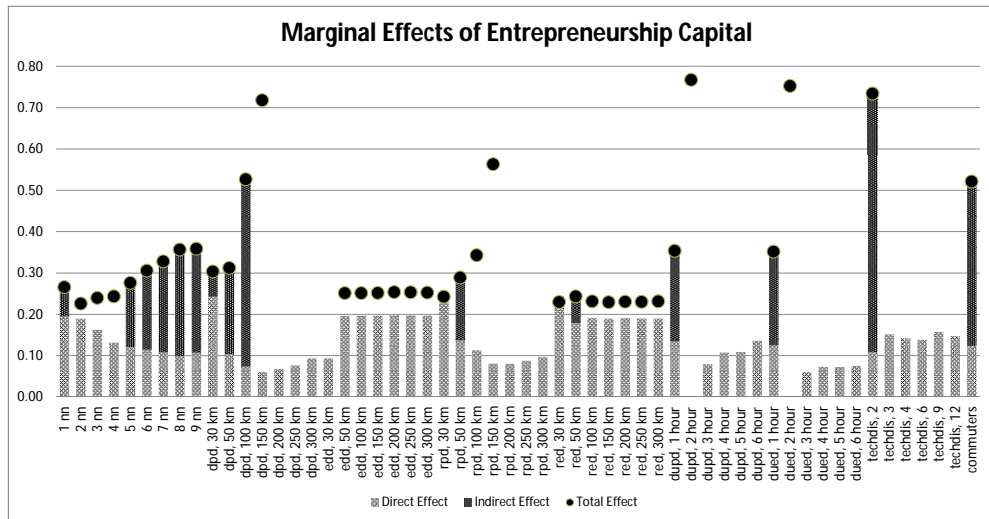
Chapter 1 Do Regions with Entrepreneurial Neighbors Perform Better? A Spatial Econometric Approach for German Regions

Figure 1.5: Coefficient comparison for different weight matrices: human capital



Notes: Coefficients are only plotted if significant, nn = nearest neighbor, dpd = direct power distance, edd = exponential direct distance, rpd = road power distance, red = road exponential distance, dupd = duration power distance, dued = duration exponential distance, techdis = technological distance

Figure 1.6: Coefficient comparison for different weight matrices: entrepreneurship capital



Notes: Coefficients are only plotted if significant, nn = nearest neighbor, dpd = direct power distance, edd = exponential direct distance, rpd = road power distance, red = road exponential distance, dupd = duration power distance, dued = duration exponential distance, techdis = technological distance

1.A.2 Tables

Table 1.4: Data description

Variable	Description	Year	Source
Output	Gross value added at basic prices	2008	Eurostat
Physical capital	Investment in manufacturing and mining, Perpetual inventory method, $\delta = 5\%$	1995-2008	German Statistical Office (Destatis)
Human capital	Share of employees with technical college or university degree in working age population	2008	German Statistical Office (Destatis)
Entrepreneurship capital	Start-up rates in the following areas: Cutting-edge technology manufacturing, High-technology manufacturing, Technology-intense services, Skill-intense services	1997-2004	Mannheim Enterprise Panel (MUP) (ZEW - Center for European Economic Research)

Chapter 2

Self-Employment and Economic Performance - A Geographically Weighted Regression Approach for European Regions

2.1 Introduction

The literature on entrepreneurship and its effect on economic development is quite scarce for a European regional setting. The main reason is that the only data available at the regional scale is the self-employment rate. Acs and Szerb (2009) argue that the conventional measure of entrepreneurship, i.e. the self-employment rate, is not appropriate. Sanandaji (2010) mentions that self-employment also includes "construction workers, shop owners, taxi and truck drivers, gardeners, plumbers, fast food vendors, hair-dressers" (Sanandaji, 2010, p.1) and so on. Those entrepreneurs are generally not seen as entrepreneurs in the Schumpeterian sense (Schumpeter, 1934), where entrepreneurs are the source of innovative activity. However, the theoretical expected positive effect of entrepreneurship on economic development comes from the innovative nature of entrepreneurship (Audretsch and Keilbach, 2004). Unfortunately, using the self-employment rate means that one is unable to distinguish between necessity and opportunity entrepreneurship. Following Acs and Varga (2005), opportunity entrepreneurs start their own business because they pursue an opportunity, while necessity entrepreneurs

start their own business because it is the best, but not the preferred, option available.

In the literature it is argued that the level of the self-employment rate could be used to measure regional entrepreneurship culture (Fritsch and Wyrwich, 2012). Entrepreneurship culture is also known as entrepreneurship capital and was introduced by Audretsch and Keilbach (2004). The authors state that entrepreneurship capital are all factors conducive to the creation of new business. In detail the authors mean by entrepreneurship capital "aspects such as a high endowment with individuals willing to take the risk of starting up a new business. It also implies the existence of a regional milieu that encourages start-up activities such as an innovative milieu, the existence of formal and informal networks, but also the general social acceptance of entrepreneurial activity and the activity of bankers and venture capital agents willing to share risks and benefits involved" (Audretsch and Keilbach, 2004, p. 951). Sorenson and Audia (2000) explain that observing successful entrepreneurs with a similar background may increase an individuals self-confidence and the likelihood to start their own business. Additionally, Fornahl (2003) develops how positive entrepreneurial examples lead to the development of an agent to become an entrepreneur. Thornton et al. (2011) claim that institutional orders may support or discourage entrepreneurial behavior. Moreover, Beugelsdijk (2007) develops that the future entrepreneurs willingness to take risk will result in increased economic dynamism, innovativeness, and thus economic growth.

Entrepreneurship culture, defined in this way, should encourage opportunity entrepreneurship and thereby exert a positive effect on economic development. This latent variable should clearly manifest itself in a high self-employment rate.

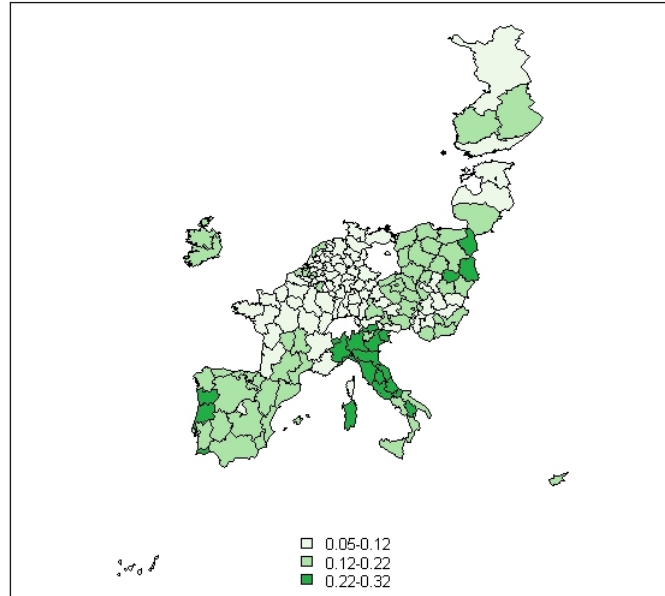
However, a high self-employment rate could also mirror a lack in wage-employment opportunities (Thurik et al., 2008). This in turn implies that a large part of the self-employed people are necessity entrepreneurs. When there are limited employment possibilities, the opportunity costs of starting

a new business decrease. Thurik et al. (2008) state that if opportunity costs decrease, people will start their own business even if they not possess the entrepreneurial talent, the knowledge, and innovative ideas necessary to start and sustain a new firm. Faggio and Silva (2012) find that there is no correlation between self-employment, business creation, and innovation in rural areas of Great Britain, and that this is related to a lack of employment opportunities. These findings suggest that the effect of self-employment on economic development could as well be insignificant. Empirically, even negative effects of self-employment on economic development are found (Blanchflower, 2000; van Stel et al., 2005). van Stel and Storey (2004) argue that a negative effect may appear when subsidized new entrants force established entrepreneurs out of the market and then, after expiration of the subsidization, leave themselves the market because they no longer have a competitive advantage. van Stel et al. (2005) explain that the negative effect may arise from the low human capital of opportunity entrepreneurs, as these entrepreneurs would be more productive as employees in large firms.

In sum, if the self-employment rate mirrors a positive entrepreneurial environment, which attracts innovative opportunity entrepreneurs, the effect on economic development should be positive. If a region is characterized by a lack in wage employment opportunities, the self-employment rate should be dominated by necessity entrepreneurs, and therefore no or a negative effect on economic development should be expected.

In our study we are interested to what extent spatial heterogeneity in the effect of the self-employment rate on economic development is prevalent at the European NUTS-2 (Nomenclature of Territorial Units for Statistics) level. Given the large variation in the self-employment rate across European regions (Figure 2.1) spatial heterogeneity is likely to occur. If the effect of self-employment on economic output in European NUTS-2 regions depends on the location, the underlying process is called to be spatially non-stationary. To deal with spatial non-stationarity, Fotheringham et al. (2002) develop the geographically weighted regression (GWR) approach. With this method, a separate regression is estimated for each region. The sample of each regression

Figure 2.1: Self-employment per working age population, 2004



Note: missing values are blank

contains the location of interest and neighboring regions, which are weighted according to their distance from the region of interest (Brunsdon et al., 1996). In this way a separate set of coefficients for each region is obtained and these can be used to visualize the regional varying effect of self-employment on economic development.

In an attempt to find out more about the source of spatial non-stationarity of the effect of self-employment, we subsequently make use of the concept of equilibrium rate of entrepreneurship. In detail, the concept of equilibrium rate of entrepreneurship is applied to estimate a level of the self-employment rate from which point relatively more entrepreneurs are self-employed out of necessity than out of opportunity. Using this concept, we try to find out whether heterogeneity in the effect of self-employment on economic development is indeed due to different relative sizes of necessity and opportunity entrepreneurship in the self-employment rate. The equilibrium rate of entrepreneurship depends on the level of economic development. According

to this concept, each level of economic development implies a certain lack of wage employment opportunities that will determine the equilibrium rate of entrepreneurship. Thus, if the self-employment rate is above the equilibrium rate there is apparently a bigger lack of wage-employment opportunities than what the level of economic development suggests. This concept of equilibrium rate of entrepreneurship is used, as it is, to our knowledge, the only attempt in the literature to estimate a level of the self-employment rate from which point the rate can be interpreted as being dominated by necessity entrepreneurs. The concept is explained in more detail in section 2.4. We expect regions having a self-employment rate above the equilibrium rate to exert no or a negative effect on economic development. Furthermore, we expect regions having a self-employment rate below the equilibrium rate to exert a positive effect on economic development, as such a rate should not be dominated by necessity entrepreneurs. Compared to rates above the equilibrium rate, rates below should mirror relatively more entrepreneurial talent and more innovative ideas. For those regions the self-employment rate could be used to measure entrepreneurship culture.

To the best of our knowledge, we are the first applying GWR to visualize the heterogeneous effects of self-employment on economic development on a regional level. Furthermore, this paper is the first one that tries to shed light on the question as to why there is spatial heterogeneity by estimating a level of the self-employment rate from which point the rate can be interpreted as being dominated by necessity entrepreneurship.

Applying GWR, we find a significant positive effect of self-employment on economic development for parts of Austria, Germany, and Italy, as well as for Estonia, Finland, Latvia and the Netherlands. Significant negative effects are found for parts of France, Portugal, and Spain. The GWR results show that in regions where the effect is significantly positive, the self-employment rate is, on average, smaller than in regions where the effect is significant negative. Using the concept of equilibrium rate of entrepreneurship, we find that in regions where the self-employment rate is below the equilibrium rate it has a

positive effect on economic development. In regions where the self-employment rate is above the equilibrium rate, the effect of self-employment on economic development is negative. We see this as evidence, that self-employment rates above the equilibrium rate can indeed be interpreted as being dominated by entrepreneurs out of necessity.

The paper is organized as follows: section 2.2 explains in more detail necessity and opportunity entrepreneurship. In section 2.3 we describe the model, and the data. Moreover, this section explains the geographically weighted regression theoretically and provides first empirical results. Section 2.4 introduces the concept of equilibrium rate of entrepreneurship and presents the final empirical results. Section 2.5 concludes.

2.2 Necessity and Opportunity Entrepreneurship

According to the Eurostat's concept of self-employment, "self-employed persons are defined as persons who are the sole owners, or joint owners, of the unincorporated enterprises in which they work" (European Commission and Eurostat, 1999, p. 38). Thus, the self-employment rate does not allow us to distinguish between necessity and opportunity entrepreneurship. Following the empirical findings by Acs and Varga (2005), necessity entrepreneurship does not result in technological change, while opportunity entrepreneurship does.

Unemployment is often found to result in necessity entrepreneurship. Dejardin and Carree (2011), for example, find that people who decide to start their own business out of unemployment choose industries like shoe stores, flower shops and fast food, which have relatively low entry barriers. Furthermore, Pfeiffer and Reize (2000) find that start-ups out of unemployment have lower survival probability. This is supported by van Stel and Storey (2004), who find that in some areas of Great Britain, which are lacking in enterprises and where policies tried to increase firm formation, the effect on employment is negative. Thus,

necessity entrepreneurs cannot be considered as a source of innovative activity.¹

However, the theoretical expected positive effect of entrepreneurship on economic development comes from the innovative nature of entrepreneurship (Audretsch and Keilbach, 2004): New businesses increase competition and force established firms to be more efficient, innovative and thus more competitive. Moreover, new firms produce variations of existing products and lead thereby to a greater diversity. Product diversity may be stimulating for economic development as it favors follow up innovations. Finally, new firms constitute an important link between knowledge creation and knowledge commercialization. It is not just the creation of knowledge that generates economic output, but rather when knowledge is commercialized. According to the Knowledge Spillover Theory of Entrepreneurship (Acs et al., 2009), a new firm is a vehicle through which knowledge spills over from the source of knowledge production, i.e. an incumbent firm or university, into the economy where it becomes economically relevant knowledge and generates economic output (Braunerhjelm et al., 2010). In a nutshell, Holcombe (2006) states "[Economic] progress occurs because of innovations introduced into the economy, and innovations are the result of entrepreneurship" (Holcombe, 2006, p. 28).

As measures like the self-employment rate cannot distinguish between necessity and opportunity entrepreneurs, it is not surprising that most studies analyzing the effect of entrepreneurship on economic development find results that are quite heterogeneous across regions. Some studies find a positive relationship between firm birth and local economic performance, like Acs and Armington (2004) for the US, or Fölster (2000) for Swedish counties. Blanchflower (2000) finds a negative effect in OECD countries. Meanwhile, others find a conditional effect: van Stel et al. (2005) detect in their analysis of 36 countries that the effect of entrepreneurial activity on economic growth

¹This conclusion is not supported by Caliendo and Kritikos (2010), who find that unemployed persons do not only create their own business out of necessity but because they see an opportunity.

depends on the level of economic development. The effect is positive in highly developed economies and negative in developing countries. The same is found by Acs et al. (1994). Moreover, Fritsch and Schroeter (2011) find that the effect of new business formation on economic performance in West Germany depends on factors like population density, the amount of innovative activity, or the share of medium-skilled workers. Likewise, Berkowitz and DeJong (2005) find for post-Soviet Russia a positive effect of entrepreneurial activity on economic growth, depending on initial conditions and policy reforms. Finally, Li et al. (2011) find that there are different relationships between business formation and economic development across metropolitan and non-metropolitan counties in the US.

2.3 Geographically Weighted Regression

2.3.1 Description of the Model when Assuming Homogeneous Effects Across Space

We follow Audretsch and Keilbach (2004) and consider a neoclassical production function, which not only includes the standard variables physical capital (K), human capital (H), and labor (L) as explanatory variables but also self-employment (E):

$$Y = F(K, H, E, L). \quad (2.1)$$

The variables are divided by labor (L) so we work with productivity expressions, y , k , h , e . We use the Cobb-Douglas specification of the production function for our analysis of the effect of self-employment on economic development:

$$y_{it} = a_i k_{it}^{\alpha_k} h_{it}^{\alpha_h} e_{it}^{\alpha_e}, \quad (2.2)$$

where $i = 1, \dots, n$ denotes regions, $t = 1, \dots, T$ denotes time, and a represents the state of the technology. Taking logs, the equation we are going to estimate in a first step, without assuming heterogeneous effects across space, is:

$$\ln y_{it} = \ln a_i + \alpha_k \ln k_{it-1} + \alpha_h \ln h_{it-1} + \alpha_e \ln e_{it-1} + \epsilon_i. \quad (2.3)$$

The dependent variable economic output, y , is measured by gross value added at basic prices; physical capital, k , is calculated with the perpetual inventory method using data on gross fixed capital formation.² Human capital, h , is measured by patent applications. Another measure of human capital, namely the share of employees with technical college or university degree, turns out to be insignificant. Self-employment, e , is included as defined by the European Commission and Eurostat (1999). All variables are divided by economically active population. Moreover, population density is included in the estimation of equation 2.3 as a control variable. All variables are available on a yearly basis. An overview on the variables and sources can be found in Table 2.2. We estimate an unbalanced panel for 178 regions from 18 European countries for the period 1999-2005. We use the fixed effects estimator to account for region specific effects. Time dummies are included to capture effects that hit all regions at the same time. To account for possible endogeneity problems the explanatory variables are lagged by one period. Given the relative short time span of 7 periods no longer time lags are considered. The regions used in the estimation are plotted in Figure 2.1.

2.3.2 First Empirical Results when Assuming Homogeneous Effects

The results of the fixed effects panel estimation can be found in Table 2.1, column two, 'Baseline'. Accordingly the self-employment rate has no effect

²The capital stock K_t is the sum of gross fixed capital formation in t and the depreciated capital stock from period $t-1$ ($K_t = I_t + (1-\delta)K_{t-1}$). The initial capital stock for the year 1995 is calculated as: $K_0 = \frac{1}{T-1} \sum_{t=1}^T \frac{K_t}{Y_t} Y_0$. We assume a depreciation rate, δ , of ten percent.

on economic output.³ Thus, for the European case one could argue that self-employment does not well represent entrepreneurship culture, as the expected positive effect is found to be insignificant. But one should be cautious because it may be the case that this coefficient only represent an average of local coefficients. In the presence of spatial non-stationarity the global coefficient provided by our panel estimation may be misleading locally.

2.3.3 Description of the Model when Assuming Heterogeneous Effects Across Space

To analyze whether our model is spatially non-stationary, geographically weighted regression is applied. Fotheringham et al. (2002) argue that social processes are often non-stationary over space. In this case global values could be misleading locally as they are simply spatial averages. The main reason for spatial non-stationarity mentioned by the authors are intrinsically different relationships across space due to spatial variations in attitudes, preferences, administrative, political, or contextual issues. This holds as well in our application when it comes to economic development (Partridge et al., 2008; Tabellini, 2010). Thus, if the effect of self-employment on economic development varies spatially, assuming a global model will deliver misleading results. GWR is a technique that allows local variations in the coefficients. This means that the estimated coefficients are specific to a region i . GWR is mainly used for the cross-section. This means that GWR results provide estimates for a specific region at a given moment in time. Geographically weighted panel regressions (GWPR) are in a very early stage of development (Yu, 2011). That is why we have to switch to the cross section to find out more about spatial non-stationarity of the model. In the estimation, we use data for 2005 as it is the largest most recent cross-section available. GWR estimation results for other years than 2005 are similar and will be provided upon request. GWR applied

³As all variables are in log form the coefficient of 0.13 for physical capital implies that a ten percent increase in the variable will increase economic output by about one percent.

to our model in a cross section takes the following form:

$$y_i = \alpha_{i0} + \alpha_{ik}k_i + \alpha_{ih}h_i + \alpha_{ie}e_i + \epsilon_i, \quad (2.4)$$

where α_{ie} is, for example, the coefficient of the self-employment rate in region i . With these coefficients we can create a map visualizing the locally varying effect of self-employment on economic development.

When using this concept it is assumed that parameters exhibit a certain degree of spatial consistency. That means that the parameter of nearby regions should be similar. This assumption is used in the estimation where different emphases is placed on different observations. In the estimation of the parameters in region i , only a subset of the full sample, those in regional proximity to region i , is used. For the next region j , which is a neighbor to region i , a similar but not identical subset of the sample is used, and so on. This approach is in contrast to the global model where the estimation is conducted using the full sample.

Two questions arise at this stage of the analysis. First, is it reasonable to assume, for our setting, that parameters of nearby regions should be similar, even across country borders? And second, is the spatial heterogeneity problem solved when accounting for fixed effects in the panel regression? Regarding the first question: The assumption of spatial heterogeneity implies that one intrinsically believes that space and locality matter for the economic development process. Every region has its own cultural history, its own attitudes, or even unique political conditions. This is true within and across borders. Even if every region is unique, there are also commonalities of regions that are physically close to one another. Commonalities are, for example, being member of the same country, speaking the same language, having the same physical locality and the same relevant market. Physical locality means that it makes a difference whether a region is situated in the center of Europe, with well developed infrastructure and low transportation costs, or whether a region is on the periphery (Puga, 2002). It makes a difference whether a region is surrounded by other regions being in the same

trade agreement like the European Union, or whether a region is close to the border of the European Union and, thus, partly surrounded by non-member regions. Furthermore, a firm's relevant market with the relevant demand, the relevant labor supply, and relevant knowledge is determined by space in general and not by country or NUTS region. Even if the literature survey by Niebuhr and Stiller (2004) shows that there is a border effect for European countries, which means that firms tend to sell mainly to their local market, this border effect decreases. The empirical results found by Breinlich (2006) prove that the trade reducing effect of borders and language within European regions decreases with time. Moreover, Puga (2002) finds that the European labor market has a strong geographical component, even after controlling for national and regional characteristics. In addition, Bottazzi and Peri (2003) find knowledge spillovers in Europe even after controlling for country and border effects. Finally, Rodríguez-Pose and Crescenzi (2008) find that knowledge spillovers in European regions are affected by distance decay effects. That implies that regions close to one another may have the same pool of knowledge and, therefore, a similar economic development process. However, as distance increases, the knowledge pool is no longer the same because knowledge spillovers are regionally bounded. As space matters, it is reasonable to assume that regions that are geographically in close distance should have similar coefficients.

Regarding the second question: the reasoning above implies that if we find spatial non-stationarity in the cross-section it should also be present in the panel regression, even if we account for region specific fixed effects. In GWR we assume that the underlying process is, to a large extent, area dependent, where, by area, we mean something larger than a region. And this cannot be captured by the fixed effects. So if we find non-stationarity in the cross-section it should also be present in the panel. In case we find non-stationarity in the cross-section, the results presented in Table 2.1, column two, '*Baseline*' are misleading. We then try to shed light on the non-stationarity problem using the concept of equilibrium rate of entrepreneurship.

2.3.4 GWR Estimation

Two problems arise with GWR. First, if the subset of the full sample is too small, standard errors will be high. Second, if the subsample is too large, coefficients will be biased because they drift across space. This problem is similar to the one we have with the global model. If the process is spatially non-stationary, a regression with a large subsample will result in estimates that are spatial averages. To overcome these problems a weighted calibration is used. Observations in close spatial proximity to region i have a larger influence in the estimation of the parameters for region i than those further away. That is why those observations have a larger weight in the sample than observations from regions further away. This weighted calibration will allow a sufficiently large subsample to overcome the problem of large standard errors, and it reduces the drift bias because more influence is attributed to the observations closer to i . This implies that the weighting of an observation is not constant but varies with i . Region j has a large weight in the estimation of region i if they are close to each other, and the weight of region j in the estimation of region l might be small if the regions are separated by a larger distance. The coefficients for a specific region i are estimated like this:

$$\hat{\alpha}_i = (X^T W_i X)^{-1} X^T W_i y, \quad (2.5)$$

where W_i is the spatial weighting matrix of region $i = 1, \dots, n$:

$$W_i = \begin{pmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{pmatrix}. \quad (2.6)$$

The diagonal elements of the individual weight matrix, w_{in} , determine the strength of the interaction between regions i and n .⁴ Every single region i has

⁴GWR is in contrast to simple ordinary least squares, $\hat{\alpha} = (X^T X)^{-1} X^T y$, where the diagonal elements of W_i in equation 2.5 are equal to one. Furthermore, this is in contrast to weighted least squares, $\hat{\alpha} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} y$, with Ω the variance-covariance matrix of the error term:

a different weight matrix. In the next section it is explained how the individual elements of the weight matrix, w_{in} , are determined.

2.3.4.1 Spatial Weighting Function

The question at this point is how the observations should be weighted. For this analysis the weighting functions that are most often used in the literature, namely the Gaussian and the bi-square kernel (Shearmur et al., 2007; Breitenecker and Harms, 2010; Müller, 2012) are applied. Using the Gaussian kernel the weighting of data will decrease according to a Gaussian curve as the distance between i and j , d_{ij} , increases. Up to bandwidth b the observations have a weight of at least 0.5.

$$w_{ij} = e^{-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2} \quad (2.7)$$

The bi-square kernel is a continuous, near-Gaussian weighting function up to bandwidth b , beyond b the weights are set to zero.

$$w_{ij} = \left(1 - \left(\frac{d_{ij}^2}{b^2}\right)\right)^2, \text{ if } d_{ij} < b, \text{ else } w_{ij} = 0 \quad (2.8)$$

Fotheringham (2009) states that the GWR results are relatively insensitive to the choice of the weighting function, but they are not insensitive to the choice of the bandwidth, b . As the density of regions in our dataset varies we cannot use just one bandwidth. A fixed bandwidth of for example 800 km is too small for the estimation of coefficients in Finland, because there are few regions and, accordingly, few data points in close proximity. The most northern Finish region Manner-Suomi would have only 3 neighbors, and the regions Southern and Eastern Ireland only six. Such a small sample would result in

$$\Omega^{-1} = \begin{pmatrix} \omega_1^{-1} & 0 & \dots & 0 \\ 0 & \omega_2^{-1} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \omega_n^{-1} \end{pmatrix}. \text{ In weighted least squares the weighting matrix does}$$

not vary with i . Moreover, the idea behind weighted least squares is to give less weight to observations with a high error variance, and not, as in GWR, to observations which are in larger distance.

large standard errors. Similarly, this bandwidth is too large for places like Austria, where the density of regions is much higher. The region Tirol would have 129 neighboring regions within a distance of 800 km. Such a large sample could result in serious drift bias. That is why, for our dataset an adaptive kernel is most appropriate. Adaptive kernel means that a fixed proportion of all observation is included in the estimation, for example 20 percent of all regions. Such a kernel is smaller in regions where the density of observations is high (like in Austrian regions) and larger in regions where the density is low (like in Finish regions). While the advantage of an adaptive kernel over a fixed kernel is obvious for regions with a high density of observations, the coefficients of regions with a low density of observations are likely to be drift biased, as they are also influenced by observations of regions which are in large distance.

2.3.4.2 Cross-Validation

The cross-validation (CV) method is used to find the optimal bandwidth. In the adaptive case the optimal bandwidth is not a certain number of kilometers, but a proportion of observations between 0 and 1.⁵ CV will allow to create the optimal weighting scheme for our estimation of equation 2.4:

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2, \quad (2.9)$$

where $\hat{y}_{\neq i}(b)$ is the fitted value of y_i , and was estimated without the observation i . In cross-validation the data is split into two segments. One is used to train the model and the other is used to validate the model. We are interested in the optimal weighting scheme, that means what proportion of the neighboring regions should be used. We start for example with a proportion of 0.3 of the nearest neighbors of region i . These observations without i are used to estimate equation 2.4. The coefficients of the so trained model are validated using observation i . $\hat{y}_{\neq i}(b)$ is the result of the validation and is compared to the actual value y_i . This is done for all regions $i = 1, \dots, n$. The sum of the

⁵Estimations are conducted in R 2.13.1 with the package spgwr (Bivand and Yu, 2011).

deviations is the CV-score for a bandwidth of 0.3. The procedure is repeated for all bandwidths between 0 and 1. The bandwidth with the lowest CV-score is used in the respective Gauss or bi-square weighting scheme.

For our estimation a proportion of 0.23 should be used with the bi-square kernel weighting function and a proportion of 0.04 should be used with the Gauss weighting function. The large difference can be explained with the construction of the weighting function. The bi-square weighting function only uses the observations up to the proportion of 0.23, observations beyond this bandwidth are set to zero. In the Gauss weighting function all observations are included. The bandwidth of 0.04 defines the observations that have a weight up to 50 percent in the estimation. Observations beyond this point have smaller weights.

2.3.4.3 Tests on Spatial Heterogeneity

As noted, in the literature it is often found that the effect of entrepreneurial activity on economic performance is heterogeneous across space. In a first step we therefore test whether the parameters in the GWR model vary significantly over space. If that is the case GWR should be preferred over OLS because it is able to explain the underlying relationship significantly better. Leung et al. (2000) propose a test on individual parameter stability over space. Test results for parameter stability of self-employment can be found in Table 2.3.⁶ It appears that in all years the coefficients of self-employment significantly vary over space. The results hold for the Gaussian and the bi-square kernel.⁷ As there is significant variation of the parameters across space, GWR can be applied in order to better understand the underlying mechanisms.

⁶Details on the test statistic can be found in Leung et al. (2000) on page 22.

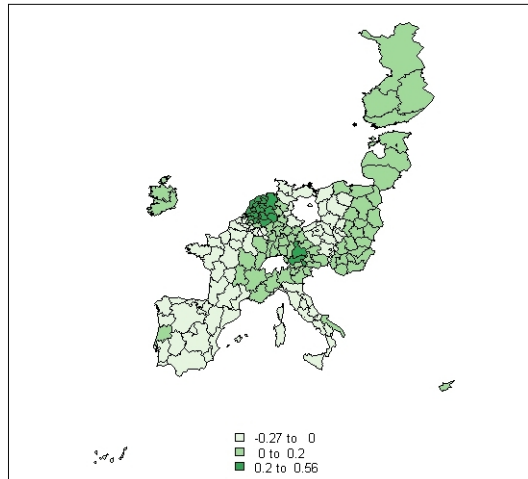
⁷Test results for the other explanatory variables reveal that their coefficients also vary across space. However, in this paper we concentrate on solving the spatial non-stationarity problem of the self-employment rate.

2.3.4.4 Empirical Results of the GWR Estimation

The estimated GWR coefficients of equation 2.4 for the two different weighting function (Gauss and bi-square) are plotted in Figure 2.2 and Figure 2.4, and the corresponding t-values in Figure 2.3 and Figure 2.5. Our results confirm Fotheringham (2009), as it appears that the results do not depend much on the choice of the weighting function. The maps show the large spatial heterogeneity of the effect of self-employment on economic performance in European regions. Taking both weighting functions into account, a significant positive effect exists for parts of Austria, Germany, and Italy, as well as for Estonia, Finland, Latvia, and the Netherlands. A significant negative effect is found for parts of France, Portugal, and Spain. In all the other regions there is no significant effect of self-employment on economic development. The GWR results show that the insignificant effect of self-employment on economic output found in the fixed effects estimation may result from the fact that positive and negative effects cancel out.

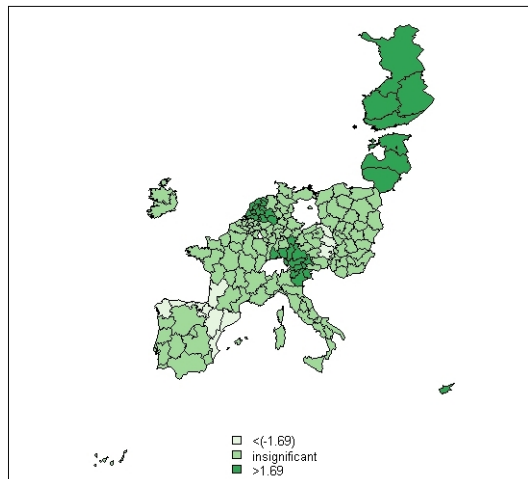
However, it is not assumed that the underlying process could best be represented using GWR, but GWR is seen as exploratory tool that points to a misspecification of the functional form. If we compare the significant positive and negative GWR coefficients with the self-employment rates, we see that while the average self-employment rate of the full sample is 14 percent, the average self-employment rate of regions having a significant positive GWR coefficient is 12.5 percent, and the average self-employment rate of regions having a significant negative GWR coefficient is 15.9 percent. This gives an initial insight that a comparably high self-employment rate may be dominated by necessity entrepreneurs. In the next step, we try to shed light on the spatial non-stationarity problem of self-employment using the concept of equilibrium rate of entrepreneurship. This concept is applied in an attempt to estimate a level of the self-employment rate from which point relatively more entrepreneurs are self-employed out of necessity than out of opportunity.

Figure 2.2: Plot of GWR parameters of the self-employment rate, α_{ie} , adaptive Gaussian weighting function



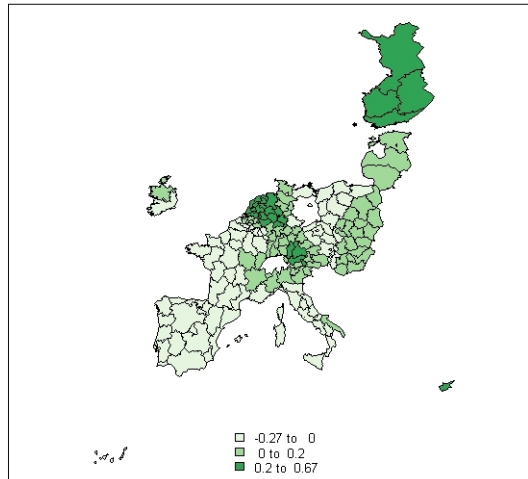
Note: Missing values are blank. Parameters of the self-employment rate vary between -0.27 and 0.56.

Figure 2.3: Plot of GWR t-values for the parameters of the self-employment rate, α_{ie} , adaptive Gaussian weighting function



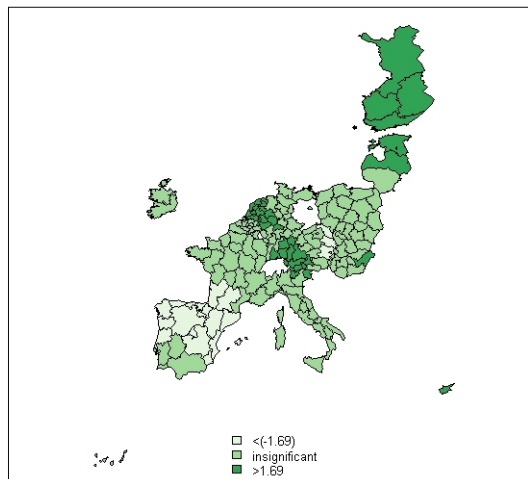
Note: Missing values are blank. The coefficients smaller than zero presented in Figure 2.2 are significant at a ten percent level in those regions, where the t-statistic is smaller than -1.69. The coefficients larger than zero presented in Figure 2.2 are significant at a ten percent level in those regions, where the t-statistic is larger than 1.69. The coefficients presented in Figure 2.2 are insignificant for t-values between -1.69 and +1.69.

Figure 2.4: Plot of GWR parameters of the self-employment rate, α_{ie} , adaptive bi-square weighting function



Note: Missing values are blank. Parameters of the self-employment rate vary between -0.27 and 0.67.

Figure 2.5: Plot of GWR t-values for the parameters of the self-employment rate, α_{ie} , adaptive bi-square weighting function



Note: Missing values are blank. The coefficients smaller than zero presented in Figure 2.4 are significant at a ten percent level in those regions, where the t-statistic is smaller than -1.69. The coefficients larger than zero presented in Figure 2.4 are significant at a ten percent level in those regions, where the t-statistic is larger than 1.69. The coefficients presented in Figure 2.4 are insignificant for t-values between -1.69 and +1.69.

2.4 The Equilibrium Rate of Entrepreneurship

Following Carree et al. (2002), the lack of wage-employment opportunities depends on the stage of economic development, as every stage has different demand conditions. The authors argue that those different demand conditions result in different rates of self-employment. In detail, Carree et al. (2002) and Bosma et al. (2005) distinguish between three stages of economic development. In the first stage per capita income is relatively low and, therefore, demand for goods and services is low as well. As a consequence large firms do not exist because they could not benefit from economies of scale and scope. A high percentage of the population in economies at this stage of development are self-employed, because many alternatives do not exist. At this stage entrepreneurial activity is negatively related to economic development (Acs et al., 2008). In the second stage per capita income is higher. This allows firms to benefit from economies of scale and scope because now there is increased demand. At this stage there are more opportunities to become an employee. Furthermore, as employees earnings increase Lucas Jr. (1978) states that this "raises the opportunity cost of managing relative to the return" (Lucas Jr., 1978, p. 518). This is why at this stage of economic development there are fewer self-employed individuals. Following Bosma et al. (2005), at a third stage incomes are higher and allow for a realization of individual preferences. That means that there is higher demand for variety and, therefore, more space for entrepreneurial ideas (Verheul et al., 2002). At this stage the rate of self-employment rises and, thus, entrepreneurial activity is again positively related to economic development. These three stages suggest a U-shaped equilibrium rate of entrepreneurship. However, Carree et al. (2007) empirically find for 23 OECD countries that not only is a U-shape consistent with the data but also an L-shape. An L-shape would imply that there is no upswing but a stabilization of the equilibrium rate as the economy develops (Wennekers et al., 2010). Moreover, Wennekers et al. (2010) state that the upswing in the U-shaped relationship is due to an increased number of opportunity entrepreneurs, because of a growing need for

independence at higher levels of economic development.

This theory implies that if the self-employment rate is above the equilibrium rate there is apparently a stronger lack of wage-employment opportunities than what the level of economic development suggests. Due to lower opportunity costs more people decide to start their own business even if they do not have the qualifications and capabilities necessary to be a successful entrepreneur. We expect the effect of self-employment on economic development to be insignificant or negative in regions having a self-employment rate above the equilibrium level. The way we estimate the different equilibrium rates is in line with Carree et al. (2002).

2.4.1 The Model

Carree et al. (2002) suggest four different specifications of the equilibrium rate of entrepreneurship, E^* , two U-shaped and two L-shaped versions:

The quadratic U-shape:

$$E_{i,t}^* = \alpha + \beta \text{YCAP}_{it} + \gamma \text{YCAP}_{it}^2, \quad (2.10)$$

the log quadratic U-shape:

$$E_{i,t}^* = \alpha + \beta \ln(\text{YCAP}_{it} + 1) + \gamma (\ln(\text{YCAP}_{it} + 1))^2, \quad (2.11)$$

the inverse L-shape:

$$E_{i,t}^* = \alpha - \beta \frac{\text{YCAP}_{it}}{\text{YCAP}_{it} + 1}, \quad (2.12)$$

and the log inverse L-shape:

$$E_{i,t}^* = \alpha - \beta \frac{\ln(\text{YCAP}_{it} + 1)}{\ln(\text{YCAP}_{it} + 1) + 1}. \quad (2.13)$$

Following the theoretical reasoning mentioned above, in every specification the equilibrium rate of entrepreneurship is a function of per capita income,

YCAP. The two U-shaped functional forms capture the above explained drop and subsequent rise in entrepreneurship as per capita income increases. The two L-shaped functional forms capture the stabilization of the equilibrium rate of entrepreneurship as per capita income increases. To estimate the equilibrium rate of entrepreneurship, the authors use the following equation, which explains changes in entrepreneurship in the following way:⁸

$$\Delta_2 E_{it} = b_1 (E_{i,t-2}^* - E_{i,t-2}) + b_2 (U_{i,t-2} - \bar{U}) + b_3 (LIQ_{i,t-2} - \bar{LIQ}) + \epsilon_{it}. \quad (2.14)$$

As this concept is applied in an attempt to shed light on the spatial non-stationarity problem of self-employment, the coefficients are estimated in a panel regression and do not vary with i . Changes in entrepreneurship are explained by deviations of entrepreneurship, E , from the equilibrium rate, E^* . If entrepreneurship lies below the equilibrium rate the rate of business ownership is expected to rise. Moreover, if the unemployment rate, U , is above the average unemployment rate over i and t , \bar{U} , the rate of business ownership is expected to rise.⁹ Finally, the deviation of the labor income share, LIQ , to the average labor income share of the sample, \bar{LIQ} , is used as explanatory variable. Labor income share is defined as the share of labor income in national income and tries to capture the earnings differentials between expected profits of entrepreneurs and employees. If this share is relatively high, i.e. labor income is a large part of national income, expected capital and entrepreneurship income are low. In this case it is less likely that a person starts their own business. An overview of the variables and sources can be found in Table 2.2. If one of the equations 2.10 to 2.13 is substituted in equation 2.14 we get the following equations:

⁸Carree et al. (2002) use a different lag specification than we do. In detail they use $E_{i,t-4}^*$, $E_{i,t-4}$, $U_{i,t-6}$, and $LIQ_{i,t-6}$. They explain that mental preparation for starting a new business needs up to six years. As our time series are not long enough, we cannot use the same lag structure. Moreover, we are not fully convinced, that mental preparation for starting a new business needs that much time.

⁹However, the OECD (2000) finds that "only a very small proportion of unemployed people find employment through self-employment" (OECD, 2000, p. 157).

For the quadratic U-shape:

$$\begin{aligned}\Delta_2 E_{it} = & a_0 - b_1 E_{i,t-2} + b_2 U_{i,t-2} + b_3 LIQ_{i,t-2} + a_4 YCAP_{it-2} \\ & + a_5 (YCAP_{it-2})^2 + \epsilon_{it},\end{aligned}\quad (2.15)$$

for the log quadratic U-shape:

$$\begin{aligned}\Delta_2 E_{it} = & a_0 - b_1 E_{i,t-2} + b_2 U_{i,t-2} + b_3 LIQ_{i,t-2} \\ & + a_4 \ln(YCAP_{it-2} + 1) \\ & + a_5 (\ln(YCAP_{it-2} + 1))^2 + \epsilon_{it},\end{aligned}\quad (2.16)$$

for the inverse L-shape:

$$\begin{aligned}\Delta_2 E_{it} = & a_0 - b_1 E_{i,t-2} + b_2 U_{i,t-2} + b_3 LIQ_{i,t-2} \\ & + a_4 \frac{YCAP_{it-2}}{YCAP_{it-2} + 1} + \epsilon_{it},\end{aligned}\quad (2.17)$$

and for the log inverse L-shape:

$$\begin{aligned}\Delta_2 E_{it} = & a_0 - b_1 E_{i,t-2} + b_2 U_{i,t-2} + b_3 LIQ_{i,t-2} \\ & + a_4 \frac{\ln(YCAP_{it-2} + 1)}{\ln(YCAP_{it-2} + 1) + 1} + \epsilon_{it}.\end{aligned}\quad (2.18)$$

We can use the estimated coefficients, a_0, b_1, b_2, b_3 , and a_4 , to calculate α , β , and γ from equations 2.10, 2.11, 2.12, and 2.13:

$$\hat{\alpha} = \frac{a_0 + b_2 \bar{U} + b_3 \bar{LIQ}}{b_1}, \quad \hat{\beta} = \frac{a_4}{b_1}, \quad \text{and} \quad \hat{\gamma} = \frac{a_5}{b_1}.$$

These coefficients are then used to calculate the equilibrium rate of entrepreneurship according to equations 2.10, 2.11, 2.12, and 2.13.

2.4.2 The Estimated Equilibrium Rates of Entrepreneurship

The estimated equilibrium rates of entrepreneurship are plotted in Figure 2.7. In the same figure the actual self-employment rate in the year 2004 is plotted. It can be seen that there are regions that are quite close to the different equilibrium rates and other regions that are apart. Furthermore, it appears that the four rates are not very different from one another. Our estimated equilibrium rates are similar to the equilibrium rate estimated by Carree et al. (2002). In the estimation we want to work with percentage deviations from those equilibrium rates, as Carree et al. (2007) do. The correlation coefficients between the deviations from the equilibrium rates are close to one. Apparently we are not in the presence of a model selection problem, the percentage deviations of the different equilibrium rates are basically the same. Between a per capita income of 10,000 and 30,000, the four lines are almost the same and the overall average equilibrium self-employment rate is 12.1 percent. When calculating the percentage deviation from this constant rate and comparing it to the percentage deviations from the different equilibrium rates, we, again, get a correlation coefficient that is close to one. Apparently, when working with European regions it is not necessary to estimate a U- or L-shaped equilibrium rate. A constant rate of about 12 percent will provide the same results. For most levels of per capita income the estimated equilibrium rate by Carree et al. (2002) is also 12 percent.

There are two possible explanations for this result. First, one could argue that the present data only mirrors the second stage of economic development, namely the stage where economic development is that high that demand for goods and services allows big enterprises to exist because they can profit from economies of scale and scope. However, in our example gross value added per capita varies between 1,260 and 54,680 Euro. This window is larger than that used in the estimation of the first application of the equilibrium rate of entrepreneurship by Carree et al. (2002) and should therefore capture the downswing of the U- or L-shaped equilibrium rate.

Thus, it may be the case that the theory of a U-, or L-shaped equilibrium rate of entrepreneurship is not appropriate for a regional setting.

When we look at the regional level it is not the case that demand of one region is necessarily met by supply of this same region. That means that demand is met by supply that has its origin in different regions from the same or other countries.¹⁰ The concept of equilibrium rate of entrepreneurship states that a low level of economic development implies low demand for goods and services and therefore no opportunities for firms to profit from economies of scale and scope. Accordingly self-employment should be high for low levels of economic development. But if a firm decides to locate where wages are comparably low, i.e. in regions with a low per capita income, and serve markets in other regions where demand and per capita income is high, the theory would no longer hold. Self-employment would be low in these regions because there are more opportunities to be wage employed. This should especially be true within a country but also between countries in the European Union. This may be the reason why the U-, or L-shaped equilibrium rate is neither appropriate for our setting nor, possibly, for other regional settings. However, it may still work for countries where most demand is met nationally.

This reasoning may also explain a constant equilibrium rate of entrepreneurship. If demand can be met by supply from different regions, there is no longer an adaption of the regional self-employment rate to changing regional demand.

2.4.3 Results Using the Equilibrium Rate of Entrepreneurship

As explained, the U-, or L-shaped equilibrium rates of entrepreneurship are not appropriate for our dataset. That is why we replace the spatial non-stationary self-employment rate with the percentage deviations of self-employment from the constant equilibrium rate. Following Carree et al. (2007) we include positive and negative percentage deviations in absolute terms separately in the fixed effects panel estimation of equation 2.2. Results can be found in Table 2.1 in the third column, '*Deviations from equilibrium rate*'. Both positive

¹⁰As mentioned above, there is a positive but decreasing border effect for European regions.

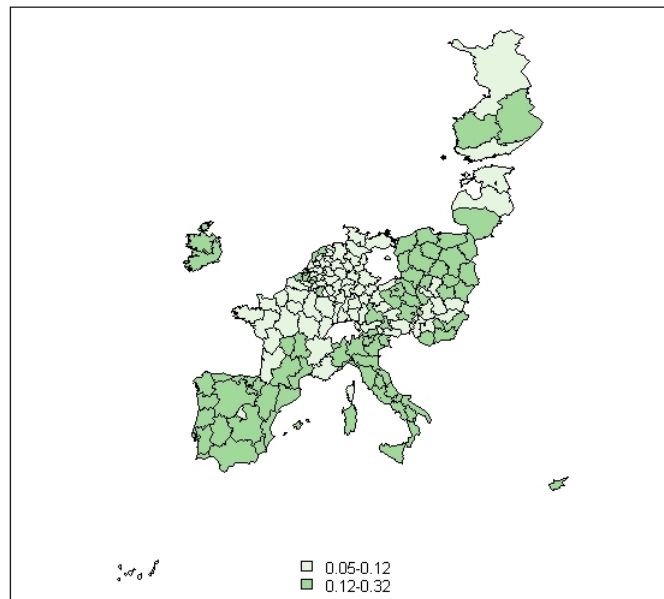
Table 2.1: Panel estimation

Dependent variable: economic output, y				
Variable	'Baseline'	'Deviations from equilibrium rate'	'Sample if rate >12%'	'Sample if rate <12%'
Physical capital $_{t-1}$	0.13*** (0.02)	0.13*** (0.02)	0.25*** (0.05)	0.08*** (0.03)
Human capital $_{t-1}$	0.01** (0.006)	0.01** (0.006)	0.00 (0.005)	0.07*** (0.02)
Self-employment $_{t-1}$	0.04 (0.03)		-0.07** (0.04)	0.08* (0.05)
Positive percentage deviation from equilibrium rate $_{t-1}$		-0.01 (0.03)		
Negative percentage deviation from equilibrium rate $_{t-1}$		-0.05 (0.06)		
Density	-1.10*** (0.18)	-1.11*** (0.18)	-1.16*** (0.18)	-0.67** (0.31)
R^2	0.53	0.53	0.71	0.41
Observations	967	967	521	446
Number of groups	178	178	109	103

Notes: Fixed effects estimation of an unbalanced panel with time dummies; annual data for the period 1999-2005; ***, **, * statistically significant at one, five, and ten percent, respectively; standard errors in parentheses. Density is included in the estimation of equation 2.3 as a control variable. To account for possible endogeneity problems the explanatory variables are lagged by one period.

and negative deviations from the equilibrium rate have a negative sign, which is not significant. That implies that a region having a self-employment rate much above the equilibrium rate does not perform significantly worse than a region having a self-employment rate only a little above the equilibrium rate. The same is true for negative deviations. Apparently the gap between actual and equilibrium rate of self-employment is not decisive. The insignificant effect of positive percentage deviations from the equilibrium rate is also found by Carree et al. (2007). However, in their analysis of OECD-countries negative percentage deviations from the equilibrium rate have a significant negative effect.

Figure 2.6: Self-employment rates above and below the constant equilibrium rate of 12 percent in 2004



Note: missing values are blank

As we interpret the constant equilibrium rate as the level of the self-employment rate from which point relatively more entrepreneurs are self-employed out of necessity than out of opportunity, in a next step we split the sample into regions having a self-employment rate above the equilibrium

rate and into regions having a rate below the equilibrium rate (Table 2.1, column four, '*Sample if rate >12%*', and five, '*Sample if rate <12%*'). In Figure 2.6 it is plotted which regions have a self-employment rate below or above the constant equilibrium rate in 2004.

In the sample where the self-employment rate is below the equilibrium rate, an increase in the self-employment rate has a significant positive effect on economic output. In the sample, where the self-employment rate is above the equilibrium rate, the effect of self-employment on economic development is negative. This confirms our initial guess, namely that the insignificant coefficient of 0.04 of self-employment in the baseline estimation (Table 2.1, column two, '*Baseline*') is just a global average. Furthermore, we see these results as evidence, that self-employment rates above the equilibrium rate can indeed be interpreted as being dominated by entrepreneurs out of necessity. In regions where the self-employment rate is just high because opportunity cost of starting a new business are low, the so created entrepreneurs are apparently not the ones that, according to Beugelsdijk (2007), increase economic dynamism, innovativeness and, ultimately, economic output.

To find out, whether this approach really solves the spatial non-stationarity problem of self-employment, we should perform the individual parameter stationarity test for the two subsamples. However, given that the number of regions in the two subsamples is comparably small for GWR, and given that the regions in the subsamples are no longer continuous in space, reasonable results of such a test cannot be expected. Thus, we cannot directly answer the question, whether the application of the concept of equilibrium rate of entrepreneurship solves the spatial non-stationarity problem of the self-employment rate. Moreover, we mentioned that the individual parameter stability test for the other explanatory variables pointed as well to a spatial non-stationarity problem. The estimation results for the two subsamples confirm the test result. The parameters of human capital (0.00 and 0.07) and physical capital (0.25 and 0.08) in the two subsamples are quite different from the coefficients estimated with the entire sample (0.01 for human capital, and

0.13 for physical capital). Thus, the results presented in Table 2.1, column four, and five, are likely to still suffer from misspecification.

However, the GWR results, together with the results of the two subsamples, shed further light on the debate as to why heterogeneous effects of self-employment on economic development are observed. Heterogeneous effects of self-employment on economic development are due to the fact that the self-employment rate does not make the distinction between opportunity and necessity entrepreneurs. The concept of equilibrium rate of entrepreneurship appears to be a suitable tool to estimate a level of the self-employment rate from which point the rate is dominated by necessity entrepreneurs.

2.5 Conclusion

The only available and comparable data on entrepreneurship at the European NUTS-2 level is the self-employment rate. However, the self-employment rate includes entrepreneurs out of opportunity and entrepreneurs out of necessity. While the effect of opportunity entrepreneurs on economic development should be positive, there should be no or a negative effect of necessity entrepreneurship. As the self-employment rate cannot distinguish between necessity and opportunity entrepreneurs it is not surprising that most studies analyzing the effect of entrepreneurship on economic development find results that are quite heterogeneous across regions.

We use a geographically weighted regression approach to find out whether the effect of self-employment on economic development is heterogeneous across European regions. We find a significant positive effect of self-employment on economic development for parts of Austria, Germany, and Italy, as well as for Finland, Estonia, Latvia, and the Netherlands. Significant negative effects are found for parts of France, Portugal, and Spain. The GWR results show that in regions where the effect is significant positive the self-employment rate is on average smaller than in regions where the effect is significant negative.

The concept of equilibrium rate of entrepreneurship is applied in an attempt

to estimate a level of the self-employment rate from which point relatively more entrepreneurs are self-employed out of necessity than out of opportunity. Necessity entrepreneurs are considered not having the qualifications and capabilities necessary to be a successful and innovative entrepreneur and, thus, are not expected to exert a positive effect on economic development.

We find that in regions where the self-employment rate is below the equilibrium rate, self-employment has a positive effect on economic development. In those regions the self-employment rate can be used to measure entrepreneurship culture. In regions where the self-employment rate is above the equilibrium rate the effect on economic development is negative. We see this as evidence, that self-employment rates above the equilibrium rate can indeed be interpreted as being dominated by entrepreneurs out of necessity.

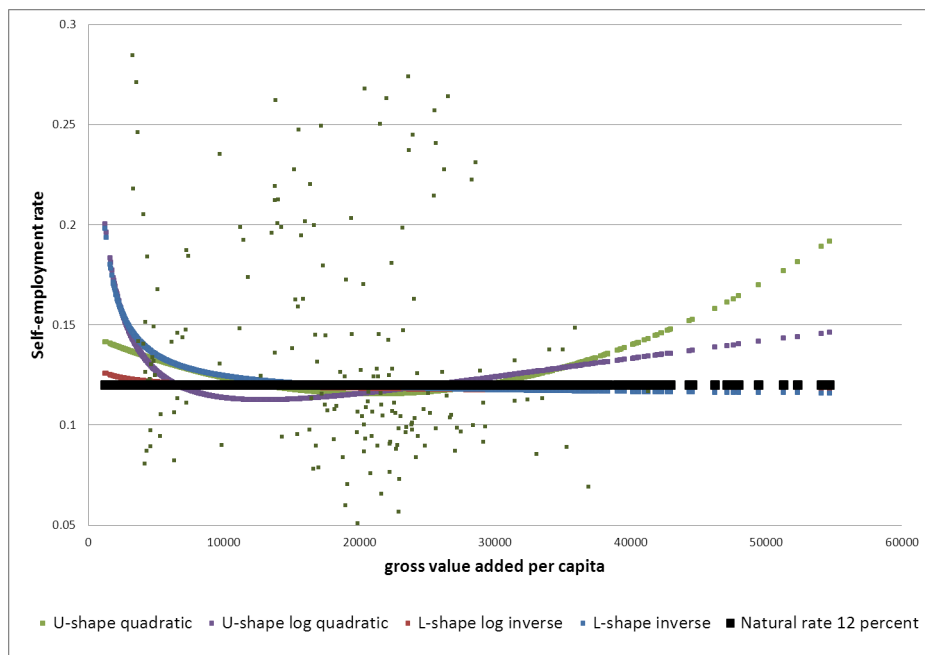
Even if we cannot say, whether the application of the concept of equilibrium rate of entrepreneurship solves the spatial non-stationarity problem, as we cannot test for it, our results shed further light to the question as to why we observe spatial heterogeneity.

The results imply that entrepreneurial research at the European regional level requires a cautious approach when using the self-employment rate. At this level the effect of the self-employment rate on economic development is insignificant, but only because positive and negative effects cancel out.

2.A Appendix

2.A.1 Figures

Figure 2.7: Plot of the estimated equilibrium rates of entrepreneurship and the actual rate of self-employment for the year 2004



2.A.2 Tables

Table 2.2: Data description

Variable	Description	Year	Source
Baseline Model			
Output	Gross value added at basic prices	1999-2005	Eurostat
Physical capital	Gross fixed capital formation, Perpetual inventory method, $\delta = 10\%$	1995-2005	Eurostat
Human capital	Patent applications at EPO	1999-2005	Eurostat
Self-employment	Self-employment rate	1999-2005	Eurostat
Equilibrium Analysis			
Per capita income	Gross value added at basic prices per capita; due to data availability issues we do not follow Carree et al. (2002) who use gross domestic product per capita. This guarantees a larger dimension in time and space.	1995-2008	Eurostat
Entrepreneurship	Self-employment rate	1999-2010	Eurostat
Unemployment	Unemployment rate	1999-2010	Eurostat
Labor income share	Calculated using total compensation of employees, total employment, number of employees and gross domestic product per capita	1999-2008, for some regions shorter time period	Eurostat

Chapter 2 Self-Employment and Economic Performance - A Geographically Weighted Regression Approach for European Regions

Table 2.3: Individual parameter stationarity test for self-employment

Year	GAUSS				BI-SQUARE			
	F-stat	df1	df2	p-value	F-stat	df1	df2	p-value
2005	1.98	37.26	139.21	0.00***	3.13	57.30	133.74	0.00***
2004	2.00	43.05	125.68	0.00***	6.02	56.11	142.21	0.00***
2003	1.88	51.37	119.67	0.00***	3.94	66.05	117.86	0.00***
2002	6.02	37.93	134.88	0.00***	7.16	51.87	114.27	0.00***
2001	1.59	17.97	109.51	0.00***	4.69	56.85	114.41	0.00***
2000	7.66	43.28	136.24	0.00***	6.49	52.49	126.91	0.00***

Note: *** statistically significant at one percent; df - degrees of freedom. F-stat is always statistically significant, in all years individual parameters of self-employment are spatially non-stationary.

Chapter 3

The Spatial Dimension of US House Price Developments

3.1 Introduction

When analyzing regional house price data, the two specific spatial aspects of regional data, namely spatial dependence and spatial heterogeneity (Anselin, 1988), need to be taken into account. This becomes clear when examining two time periods of annual house price growth rates in US metropolitan statistical areas (MSAs) plotted in Figure 3.1. The strong annual house price growth rates observed in California in 2004 are apparently transmitted to the northern coastal regions, as higher rates can be observed in those regions in 2005. This transmission of house price developments across space is what is called spatial dependence. Furthermore, the two maps suggest that house price dynamics in coastal regions are different from house price developments in the inland. This difference in the dynamics is called spatial heterogeneity. Moreover, as the house price spillovers appear to be stronger along the coastal regions it could be the case that there are differences in spatial house price spillovers across space and possibly across time, i.e. that there is heterogeneity in spatial dependence. In our analysis of regional house price developments in the US we jointly analyze these three aspects of spatial data, namely spatial dependence, spatial heterogeneity and heterogeneity in spatial dependence.

Spatial dependence and spatial heterogeneity are two well established aspects of house price developments. However, differences in spatial spillovers

across space and time have not gained much attention yet. A possible explanation of those differences in spillovers could be the so-called disposition effect (Shefrin and Statman, 1985). The disposition effect labels the phenomenon in financial markets that investors sell their winning stocks too soon and hold their losing stocks too long. Applied to the real estate market this would mean that homeowners hold their houses even if they get strong signals of declining house prices. The disposition effect implies reduced house price spillovers in times of declining house prices. Assuming incorrectly homogeneous spillovers across space and time could locally give a misleading picture of house price dynamics.

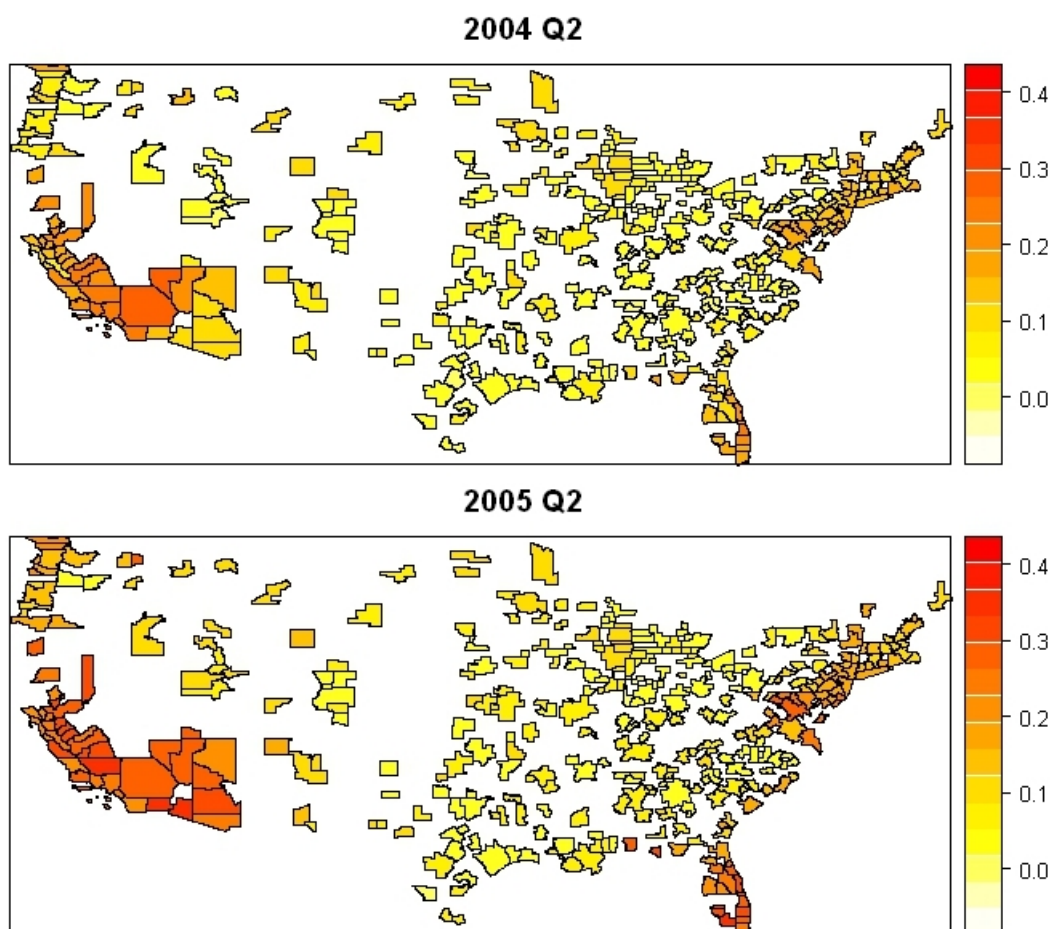
We use a panel smooth transition regression model which allows us to jointly analyze these three aspects of spatial data. González et al. (2005) develop this model in order to describe heterogeneous panels, where the coefficients can vary between regions and with time. Spatial dependence is introduced by including the spatial lag of house price developments, spatial heterogeneity is introduced by allowing the fundamentals to have heterogeneous effects across time and space, and heterogeneity in house price spillovers is introduced by allowing the spatial dependence parameter to change over time and with space. To the best of our knowledge this is the first paper providing a joint analysis of all three spatial aspects, the first paper that explicitly models heterogeneity across time and space in spatial dependence, and the first paper which tries to model the disposition effect using heterogeneity in spatial spillovers.

The results reveal that house price developments in neighboring regions spill over more in times of increasing neighboring house prices than during times of declining neighboring house prices. This is seen as evidence for the disposition effect. Heterogeneity in the effect of the fundamentals on house price dynamics is only found for population growth and building permits, but not for real per capita disposable income and the unemployment rate. The detected heterogeneity in the effect of population growth on house price developments suggest that fundamentals serve less explaining the house price developments

in times of declining house prices compared to strongly increasing house prices.

The paper is organized as follow. In section 3.2 we explain in more detail the theory and the empirical evidence of the three spatial aspects. In section 3.3 we present the econometric approach. Section 3.4 describes the available data. In section 3.5 we provide and discuss the empirical results, while section 3.6 concludes.

Figure 3.1: Annual growth rate of regional house prices, US metropolitan statistical areas



3.2 Theoretical Aspects and Empirical Evidence

Spatial dependence in house price developments is also known as the ripple effect. Accordingly, house price developments in one region cause house price movements in neighboring regions (Giussani and Hadjimatheou, 1991; Meen, 1999). Migration, equity transfer, information asymmetries and the spatial patterns in the fundamentals of house prices play a key role in the spatial spillovers of house prices (Meen, 1999). Migration or equity transfer to regions where house prices are comparably low could lead to the ripple effect by increasing demand and thereby prices. Information asymmetries may imply that new information regarding the housing market available in one area are transmitted only gradually to other sub-markets. Finally, the ripple effect could appear if variables explaining house price developments show themselves a spatial pattern.

Empirical evidence regarding spatial spillovers of house price developments is quite strong. Kuethe and Pede (2011) find in their analysis of house prices in Western United States that instate housing price forecasts can be improved by using housing prices in neighboring states. Furthermore, their results indicate that previous house prices in space and time impact current house prices. Similarly, Holly et al. (2011) find dynamic spillover effects of house prices from the neighboring regions. The diffusion of regional house prices in California counties across space is found to last up to two and a half years (Brady, 2011).

But, housing markets exhibit not only spatial dependence but also spatial heterogeneity. Following Wood (2003), one reason for spatial heterogeneity could be that some regions respond more rapidly to national economic shocks than others because their housing market is more liquid and new information is reflected more quickly in the house prices. Meen (1999) argues that heterogeneity arises because of different household behavior and household composition. Moreover, the supply of housing could be limited by planning

constraints or by geographical constraints like mountains or lakes. Thus, house prices react differently to changes in demand conditions if supply cannot adjust.

Empirical evidence for spatial heterogeneity is found by van Dijk et al. (2011), who detect the existence of two clusters of regions in the Netherlands. Regions within the cluster have the same house price dynamics, while the dynamics are different across clusters. The different clusters can be distinguished among others by the average growth rate of house prices. Furthermore, Dieleman et al. (2000) detect three clusters in their analysis of 27 metropolitan housing markets in the US, where the clusters were chosen based on the average median price and rent level. The authors find that house prices are geographical autocorrelated within the cluster.

Heterogeneity in house price spillovers over time is analyzed by de Bandt and Malik (2010) and de Bandt et al. (2010). The authors find stronger spillovers in crises times compared to normal times. A possible explanation of those differences in spillovers across time could be the so-called disposition effect (Shefrin and Statman, 1985). The disposition effect labels the phenomenon in financial markets that investors sell their winning stocks too soon and hold their losing stocks too long. Applied to the real estate market this would mean that homeowners hold their houses even if they get strong signals of declining house prices.

The prospect theory (Kahneman and Tversky, 1979), mental accounting (Thaler, 1999), and cognitive dissonance (Festinger, 1957) are concepts that may explain such behavior.

According to the prospect theory, individuals follow an S-shaped value function. Starting from a reference point, this function is concave downward above this reference point and concave upward below this reference point. This implies that individuals are risk-seeking for wealth levels below the reference point. Thus, they hold on their assets too long and, thereby take bigger risks. Genesove and Mayer (2001) find that sellers in the housing market are averse to realizing nominal losses as predicted by the prospect theory. The same is found by Engelhardt (2003).

The concept of mental accounting means that individuals group elements of their consumption and expenditures in mental accounts. They follow their personal rules in managing those accounts and react in different ways to the investments in the different accounts. When homeowners hold on their losing asset, it is because in their mental account the loss is only booked when the asset is sold and this may explain the disposition effect.

Finally, "cognitive dissonance is the mental conflict that people experience when they are presented with evidence that their beliefs or assumptions are wrong" (Shiller, 1999, p. 1314). People experiencing cognitive dissonance try to trivialize or avoid the new information, developing explanations as to why their current beliefs or assumptions should not be revised. For the housing market this would mean that homeowners avoid the information of declining house prices in neighboring regions or try to find explanations as to why this decline only applies to the neighboring region.

In sum, the disposition effect implies reduced spatial spillovers in times of declining house prices.

However, in a spatial setting the disposition effect may not only explain heterogeneity in spatial spillovers across time but also across space. This is the case as the relevant signals regarding house price developments in a given region are likely to come from neighboring regions. If house prices in neighboring regions decline, the spatial spillover is expected to be smaller compared to spillovers in regions, where house prices still increase. This implies different house price spillovers at a given moment in time because of heterogeneous house price developments across regions within the country.

In addition, heterogeneity in spatial dependence across space could arise because of different migration patterns. Kosfeld (2007), for example, finds differences in labor mobility between East and West Germany, and Molloy et al. (2011) find differences in labor mobility across US regions. Furthermore, differences in the liquidity of the housing market and, thereby, in the transmission of price information and search costs could result in different amounts of information asymmetries across regions. Kodres and Pritsker (2002) find, in their model analyzing financial contagion, that information

asymmetries may lead to stronger spillovers of financial shocks. They explain that due to information asymmetries shocks hitting only neighboring regions are mistakenly seen as shocks also hitting the region under consideration. de Bandt and Malik (2010) argue that even if housing markets are different from financial markets, financial contagion could also occur in the housing market. Thus, information asymmetries could result in heterogeneity of spatial house price spillovers across space. To our knowledge, the only paper analyzing heterogeneity in house price spillovers across space is Gray (2012). In his analysis of house price movements in England and Wales, the author finds differences in house price spillovers across space.

3.3 Econometric Approach

In the first step, we simply estimate a spatial panel fixed effects regression. This allows us to get an idea of the overall spatial spillover effect of house price developments. In a second step, the panel smooth transition regression model is estimated in order to capture the heterogeneity in spatial dependence across time and space as well as the heterogeneity in the effect of the fundamentals.¹

3.3.1 Spatial Fixed Effects Panel Estimation

Including a spatial lag of the dependent variable in our panel estimation allows us to capture the spatial spillovers of house price developments of the neighboring regions. Which regions are defined as neighbors is determined by the spatial weight matrix W_N , of dimension $N \times N$ (Anselin et al., 2008). Each element of the weight matrix, w_{ij} , determines the strength of the interaction between regions i and j . If there is no interaction between region i and j , w_{ij} is equal to zero. By convention, the diagonal elements, w_{ii} , are equal to zero. In our estimation the weights are equal to the inverse distance, $\frac{1}{d_{ij}}$, where d_{ij} is

¹Smooth transition autoregressive models were first used in time series analysis to model nonlinearity and asymmetric response (Teräsvirta, 1994).

the distance between region i and j . All regions within a distance of 330 km are going to have a positive weight in the spatial weight matrix. This distance is the smallest possible, where all regions have at least one neighbor. Furthermore, the weights are row standardized, which means that the elements of each row sum up to one. This transformation implies that W_N is no longer symmetric. For the panel case, the cross-sectional weight matrix W_N is stacked t times. Thus, we assume that the spatial weight matrix does not change over time.

$$W_{NT} = I_T \otimes W_N \quad (3.1)$$

The vector of spatially lagged dependent variables is written as:

$$Wy = W_{NT}y = (I_T \otimes W_N) y. \quad (3.2)$$

As mentioned above, regions are hit by house price development of neighboring regions with a certain time lag. Following the Akaike information criterion we include the spatially lagged dependent variable lagged by one period. Anselin et al. (2008) call this a pure space recursive model. Pure space recursive models can simply be estimated by ordinary least squares (OLS) (Lee and Yu, 2010). The general notation of the fixed effects spatial lag model we are going to estimate is:

$$y_t = \rho(I_T \otimes W_N) y_{t-1} + (\iota_t \otimes \mu) + X\beta + \nu_t, \quad (3.3)$$

where y is the dependent variable, ρ the spatial dependence parameter, and μ the region specific fixed effect. In our spatial fixed effects panel estimation, the dependent variable, real quarterly house price growth rate, hpi , is regressed on annual population growth, *population*; annual growth of the unemployment rate, *unemployment*; annual growth of real per capita disposable income, *income*; the log of building permits per population, *building permits*; the spatial lag of the dependent variable, $Whpi$; and on quarterly time dummies (Equation 3.4). We include time dummies to capture changes in the fundamentals that hit all regions at the same time, like for example changes in

the federal funds rate. Furthermore, we estimate a fixed effects regression to capture all time invariant region specific effects, μ_i . The different lags of the explanatory variables are chosen based on the Akaike information criterion.²

$$\begin{aligned} \text{hpi}_{it} = & \mu_i + \alpha \text{population}_{it-3} + \beta \text{unemployment}_{it-2} + \zeta \text{income}_{it-2} \\ & + \delta \text{building permits}_{it-1} + \rho \sum_{j=1}^N w_{ij} \text{hpi}_{jt-1} + \nu_{it}. \end{aligned} \quad (3.4)$$

The results of this estimation give an idea of the overall spatial dependence in house price dynamics and of the effect of the fundamentals, disregarding any heterogeneity across space or time.

3.3.2 Fixed Effects Panel Smooth Transition Regression Model

To include heterogeneity in the model, a non-dynamic fixed effects panel smooth transition regression model (PSTR) is estimated (González et al., 2005). This model allows the coefficients of the explanatory variables to vary between regions and with time. The coefficients change smoothly as a function of the transition variable and are, thereby, a continuous function of this transition variable. This model appears to be especially appropriate for our setting, as it allows the spatial dependence coefficient and the coefficients of the fundamentals to vary across space and time. Thereby, spatial heterogeneity in house price developments is modeled by the changing coefficients of the fundamentals; heterogeneity in house price spillovers is modeled by the changing spatial dependence coefficient. Following González et al. (2005), the PSTR model is written as follows:

$$y_{it} = \mu_i + \theta'_0 x_{it} + \theta'_1 x_{it} g(q_{it}; \gamma, c) + \nu_{it}, \quad (3.5)$$

where $g(q_{it}; \gamma, c)$ is the transition function, which is normalized to be bounded between zero and one. When the transition function is equal to zero, the coeffi-

²The results barely change when different lag specifications are used.

cient of a given explanatory variable is θ_0 ; when the transition function is equal to one, the coefficient is $\theta_0 + \theta_1$. The transition variable is q_{it} . González et al. (2005) develop the PSTR model for the logistic specification of the transition function. For the case of two extreme regimes, the logistic transition function is given by:

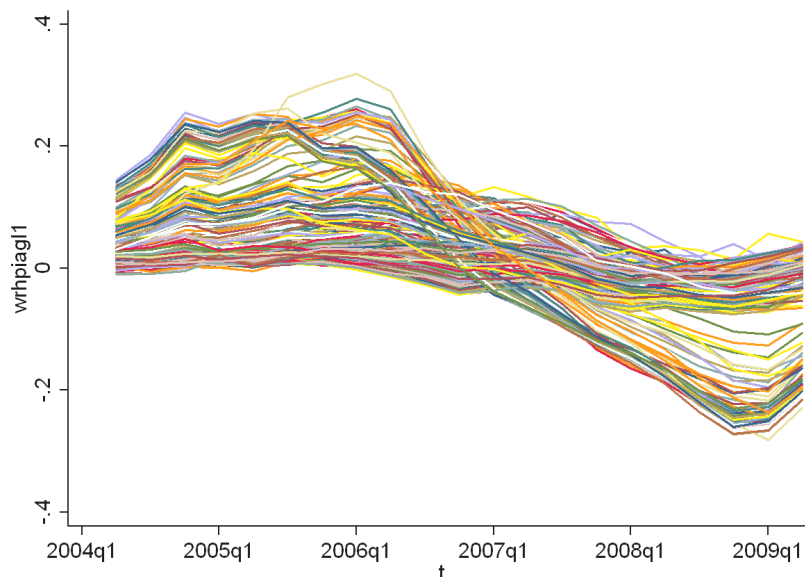
$$g(q_{it}; \gamma, c) = \frac{1}{1 + e^{-\gamma(q_{it}-c)}}, \quad (3.6)$$

where c is the location parameter, and γ is the slope of the transition function that determines the smoothness of the transition between two regimes. The higher γ , the faster is the transition between two regimes. For γ going to infinity the transition is instantaneous. For γ going to zero, the transition function becomes constant and this implies that there are no regimes at all. If the transition variable, q_{it} , is smaller than the location parameter, c , the transition function, $g(q_{it}; \gamma, c)$, tends to zero, and the coefficients tend to θ_0 . If the transition variable, q_{it} , is larger than the location parameter, c , the transition function, $g(q_{it}; \gamma, c)$, tends to one, and the coefficients tend to $\theta_0 + \theta_1$.

In the empirical application, the transition variable should capture the source of the parameter heterogeneity. A good candidate for the transition variable in our analysis is the spatially weighted house price development of the neighboring regions, $Whpiag_{t-1}$, (Figure 3.2). There are several reasons for this choice: First, de Bandt and Malik (2010) and de Bandt et al. (2010) find stronger spillovers in crises times when compared to normal times. Normal times in the housing market could be expressed by relatively normal average house price developments, while crisis times would be expressed by extreme increases or decreases in housing prices. Taking spatially weighted neighboring house price developments should be a good approximation for the overall house price development in the larger geographical region where an MSA is located.

Second, this variable seems to be a good candidate for the transition variable, as Dieleman et al. (2000) and van Dijk et al. (2011) find different house price developments within clusters, where the clusters are defined by average house

Figure 3.2: Transition variable - spatially weighted annual real house price growth rate of neighboring regions, $Whpiag_{t-1}$



price growth rates. Again, neighboring house price developments should be a good approximation of the average house price development in the larger geographical region where an MSA is located.

Third, the disposition effect for the housing market implies that a region does not react as strongly to signals of declining house prices as to signals of increasing house prices. It is reasonable to assume that those signals come from neighboring regions. Therefore, the transition variable formed by the spatially weighted house price development of the neighboring regions should be able to model the disposition effect in the housing market. In this respect, the logistic specification of the PSTR model, compared to an exponential specification, is appropriate for our setting as we expect different spillovers depending on whether we are above or below the location parameter, c . This consideration together with the fact that the estimation is conducted over a relative short time span of 21 quarters, it is reasonable to assume that there are only two extreme regimes.

In sum, the transition variable allows us to capture heterogeneity in spatial

house price spillovers and heterogeneity in the effect of the fundamentals on house price developments. However we are aware that this transition variable does not capture heterogeneity in spatial spillovers due to different migration patterns or different amounts of information asymmetries.

For estimating the parameters θ'_0 , θ'_1 , γ , and c , in a first step the region specific means need to be removed to eliminate the region specific effects, μ_i . The model can then be written as:

$$\tilde{y}_{it} = \theta' \tilde{x}_{it}(\gamma, c) + \tilde{\nu}_{it}, \quad (3.7)$$

where $\tilde{y}_{it} = y_{it} - \bar{y}_{it}$, $\tilde{x}_{it}(\gamma, c) = (x'_{it} - \bar{x}'_i, x'_{it}g(q_{it}; \gamma, c) - \bar{z}'_i(\gamma, c))'$, and $\bar{z}_i(\gamma, c) = \frac{1}{T} \sum_{t=1}^T x_{it}g(q_{it}; \gamma, c)$. This model is linear in θ given γ and c . However, the matrix of transformed explanatory variables, $\tilde{X}_{it}(\gamma, c)$, depends on c and γ and thereby needs to be recomputed at each iteration in the nonlinear least squares (NLS) estimation. c and γ are determined by applying NLS to the concentrated sum of squared errors:

$$Q^c(\gamma, c) = \sum_{i=1}^N \sum_{t=1}^T \left(\tilde{y}_{it} - \hat{\theta}(\gamma, c) \tilde{x}_{it}(\gamma, c) \right)^2, \quad (3.8)$$

where $\hat{\theta}(\gamma, c)$ is obtained from OLS of equation 3.7.

The corresponding PSTR model for our estimation of US regional house price dynamics is written as:

$$\begin{aligned} \text{hpi}_{it} &= \mu_i + \alpha_0 \text{population}_{it-3} + \beta_0 \text{unemployment}_{it-2} + \zeta_0 \text{income}_{it-2} \\ &+ \delta_0 \text{building permits}_{it-1} + \rho_0 \sum_{j=1}^N w_{ij} \text{hpi}_{jt-1} \\ &+ [\alpha_1 \text{population}_{it-3} + \beta_1 \text{unemployment}_{it-2} + \zeta_1 \text{income}_{it-2} \\ &+ \delta_1 \text{building permits}_{it-1} + \rho_1 \sum_{j=1}^N w_{ij} \text{hpi}_{jt-1}] g(\text{Whpia}_{it-1}; \gamma, c) \\ &+ \nu_{it}. \end{aligned} \quad (3.9)$$

Later on we will use the notation $\Theta_p = (\alpha_p \beta_p \zeta_p \delta_p \rho_p)$ for $p = (0, 1)$. If we assume two extreme regimes, the transition function, $g(\text{Whpiag}_{it-1}; \gamma, c)$, will tend to one for high values of the transition variable, Whpiag_{it-1} . This implies that whenever we observe high neighboring house price growth rates, the spatial spillover parameter will tend to $\rho_0 + \rho_1$, and correspondingly the other explanatory variables. Whenever we observe low or decreasing neighboring house price growth rates, the spatial spillover parameter will tend to ρ_0 , and correspondingly the other explanatory variables. In the estimation we expect ρ_0 to be smaller than $\rho_0 + \rho_1$, as the disposition effect predicts smaller house price spillovers in times of declining house prices.³

3.4 Data Description

We use data for 319 metropolitan statistical areas. Out of the existing 366 MSAs in 2012, we chose this sample due to data availability issues. Figure 3.4 plots the 319 MSAs used in the estimation. Following the United States Census Bureau, "the general concept of a metropolitan area is that of a large population nucleus, together with adjacent communities having a high degree of social and economic integration with that core" (Federal Register, 2010, p. 37246).

For the dependent variable, house prices, the Federal Housing Finance Agency all-transactions quarterly index is used. Furthermore, data on annual nominal per capita disposable income and population come from the Bureau of Economic Analysis. Annual population data is transferred to quarterly data using linear interpolation. Annual per capita disposable income is transferred to quarterly data using cubic spline. Applying cubic spline has the advantage that the resulting data is smooth in the first derivative. In order to get real values of house prices and per capita disposable income, the nominal values are divided by the consumer price index. Using real variables is in line with the literature on house price dynamics. However, this will not allow us to exactly compare our results to Genesove and Mayer (2001) and Engelhardt (2003) who

³The estimations are performed in Matlab (R2009a). We would like to thank Christophe Hurlin for kindly providing his STAR-Panel code.

find that loss aversion in the housing market depends on nominal loss. The annual consumer price index (all urban consumers) is obtained from the Bureau of Labor statistics. The regional consumer price index is not available for all MSAs used in our sample.⁴ Again, quarterly data are obtained by using cubic spline. Monthly unemployment rates from the Bureau of Labor Statistics are transferred to quarterly data by taking the quarterly averages. Population density is calculated using the population data mentioned above and the land area taken from the TIGER/Line shapefiles from the United States Census Bureau. Monthly new privately owned housing units authorized are taken from the United States Census Bureau.⁵ Quarterly data is obtained by using the sum of the monthly building permits. Building permits are divided by population. Data details can be found in Table 3.1.

As presented in equation 3.4, the quarterly growth rate of the real house price index is used in the estimation, as this time series is stationary. The explanatory variables annual growth rates of real per capita disposable income, population, and unemployment, as well as the log of building permits per population are all stationary.⁶

Unfortunately data for the effective mortgage interest rate is not available for all 319 MSAs (Mikhed and Zemčík, 2009). However, we include time dummies to capture changes in the fundamentals that hit all regions at the same time, like for example changes in the federal funds rate.

⁴Regional consumer price indexes are available for New York-Northern New Jersey-Long Island, NY-NJ-CT-PA; Philadelphia-Wilmington-Atlantic City, PA-NJ-DE-MD; Boston-Brockton-Nashua, MA-NH-ME-CT; Pittsburgh, PA; Chicago-Gary-Kenosha, IL-IN-WI; Detroit-Ann Arbor-Flint, MI; St. Louis, MO-IL; Cleveland-Akron, OH; Minneapolis-St. Paul, MN-WI; Milwaukee-Racine, WI; Cincinnati-Hamilton, OH-KY-IN; Kansas City, MO-KS; Washington-Baltimore, DC-MD-VA-WV; Dallas-Fort Worth, TX; Houston-Galveston-Brazoria, TX; Atlanta, GA; Miami-Fort Lauderdale, FL; Tampa-St. Petersburg-Clearwater, FL; Los Angeles-Riverside-Orange County, CA; San Francisco-Oakland-San Jose, CA; Seattle-Tacoma-Bremerton, WA; San Diego, CA; Portland-Salem, OR-WA; Honolulu, HI; Anchorage, AK; Phoenix-Mesa, AZ and for Denver-Boulder-Greeley, CO. For the other MSAs used in our sample the following regional consumer price indexes are used: Northeast, Midwest, South, West.

⁵We thank Konstantin A. Kholodilin for kindly providing these data.

⁶We do not include quarterly growth rates of the explanatory variables, even if the dependent variable is expressed in quarterly growth rates, because the quarterly data of real per capita disposable income and population is generated artificially from annual data.

Table 3.1: Data description

Variable	Description	Time span	Source
House price index	House prices, all transactions index	quarterly, 1986Q4-2010Q4	Federal Housing Finance Agency
Nominal per capita disposable income		annually, 1977-2009	Bureau of Economic Analysis
Population		annually, 1977-2009	Bureau of Economic Analysis
Consumer price index	Consumer price index (all urban consumers)	annually, 1984-2010	Bureau of Labor Statistics
Unemployment rate		monthly, 1990M1-2011M3	Bureau of Labor Statistics
Land area	Tiger/Line shapefile		United States Census Bureau
Building permits	New privately owned housing units	monthly, 2004M1-2012M9	United States Census Bureau

3.5 Empirical Results

To get an idea of the overall spatial dependence, disregarding any heterogeneity across space or time, we start with the spatial fixed effects panel estimation, where real quarterly house price growth rates are regressed on annual population growth, annual growth of the unemployment rate, annual growth of real per capita disposable income, the log of building permits per population, the spatial lag of the dependent variable, and on quarterly time dummies. Because data on building permits are only available over a rather short time span, the estimation is conducted for the period 2004Q2 to 2009Q2 (Table 3.2). Over this time span a balanced panel is available for 319 regions. The standard errors presented are Huber-White heteroscedasticity consistent (Huber, 1967; White, 1980). Those standard errors allow for not just heteroscedasticity in the standard errors but also for autocorrelation among observations within one cluster, where clustering takes place by MSA. However, the Huber-White standard errors do not allow for correlation among observations across clusters. This could be an inappropriate constraint as, for example, different MSAs within one state are faced by the same state specific laws. One could include state dummies to capture state effects and common shocks hitting only one state, but this is not feasible, as some MSAs cross state boundaries. Therefore, for robustification, we also estimate Driscoll-Kraay standard errors, which allow for heteroscedasticity and are robust to cross-sectional and temporal dependence as the time dimension becomes large (Driscoll and Kraay, 1998). In this way we allow for unobservable common disturbances, even across MSAs. The estimated Driscoll-Kraay standard errors are similar to the Huber-White standard errors presented in Table 3.2. The Driscoll and Kraay (1998) standard error estimates are robust as the time dimension gets large. However, the time dimension in our estimation is only 21. Therefore and because the Driscoll-Kraay standard errors are very similar to the Huber-White standard errors, we simply present the Huber-White standard errors. In the nonlinear estimation we proceed the same way.

With this spatial panel estimation we want to capture the spatial spillovers in house price developments, however, we cannot make the distinction between spatial spillover and common shocks that hit some MSAs instantaneously and some with a time lag.⁷ It could be, that a common shock hits first some regions with a very liquid housing market and reaches others with a certain time delay. Our estimation would mistake this different timing in the reaction to a common shock as spillover of house price developments. As we cannot differentiate between common shocks with region specific reaction time and spatial spillovers, we probably overestimate the coefficient of the spatially lagged dependent variable.

Table 3.2: Spatial panel estimation

Dependent variable: quarterly real house price growth rate, hpi		
Variable	FE coefficients	Huber-White standard errors
α population _{<i>t</i>-3}	0.26***	0.10
β unemployment _{<i>t</i>-2}	-0.003	0.003
ζ income _{<i>t</i>-2}	0.14***	0.02
δ building permits _{<i>t</i>-1}	0.003***	0.0007
ρ W×hpi _{<i>t</i>-1}	0.73***	0.04
R^2	0.61	
AIC	-8.3692	
Observations	6699	
Number of groups	319	
Time period	2004Q2-2009Q2	
Time dummies	yes	

Notes: Fixed effects (FE) estimation with robust standard errors to conditional heteroskedasticity of unknown form; ***, **, * statistically significant at one, five, and ten percent, respectively. Variables: *population* = annual population growth, *unemployment* = annual growth of the unemployment rate, *income* = annual real per capita disposable income growth, *building permits* = log building permits per population, W×hpi = spatially weighted quarterly real house price growth rate of neighboring regions.

The results point to a strong spillover effect of neighboring house price developments. Furthermore, the estimation results reveal a positive effect

⁷Common shocks, hitting all regions at the same time are captured by the time dummy variables.

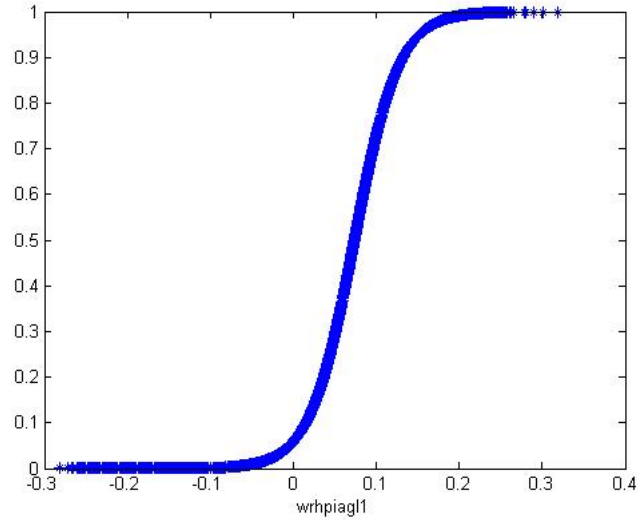
of population and real per capita disposable income growth on house price growth rates. An expected negative effect of increasing unemployment rates on house price growth rates turns out not to be significant. Furthermore, more building permits are associated with higher house price growth rates. This is an interesting result, as one could also assume that more building permits increase the supply of available housing and thereby reduce house prices. However, the coefficients assume the same strength of the spatial spillover or the same effect of the fundamentals no matter which region or time period we are looking at. These global effects could be misleading locally. The panel smooth transition regression model will help determine whether these coefficients hold for all regions in every single time period.

The results of the panel smooth transition regression in the case of two extreme regimes are presented in Table 3.3, the corresponding transition function is plotted in Figure 3.3.⁸ For almost all negative values of the transition variable, i.e. decreasing neighboring house prices, the transition function is equal to zero. This implies that the coefficients of the different explanatory variables are equal to Θ_0 in case of decreasing neighboring house prices, where $\Theta_0 = (\alpha_0 \beta_0 \zeta_0 \delta_0 \rho_0)$. Furthermore, the transition function is equal to one for very high growth rates of neighboring house prices of above 15 percent. For those high growth rates of neighboring house prices, the coefficients of the explanatory variables are equal to $\Theta_0 + \Theta_1$. The location parameter, c , is equal to 0.07, which implies that when neighboring house prices grow, on average, by 7 percent, the coefficient is equal to $\Theta_0 + 0.5 \times \Theta_1$. The slope parameter, γ , is equal to 37.49. In the literature such a value is not assumed to imply an instantaneous transition between regimes but a relatively smooth transition (Trupkin and Ibarra, 2011; Lee and Chien, 2011).

Most interestingly, the coefficient of the spatially lagged dependent variable is much smaller in case of decreasing neighboring house prices, $\rho_0=0.33$, than in case of strongly increasing neighboring house prices, $\rho_0 + \rho_1=0.6$. That

⁸The case of three extreme regimes is analyzed in the appendix.

Figure 3.3: Transition function vs. transition variable, two extreme regimes



Notes: Transition variable - spatially weighted annual real house price growth rate of neighboring regions, $Whpiag_{t-1}$

means that house price developments in neighboring regions spill over more in times of increasing neighboring house prices than during times of declining neighboring house prices. This is interpreted as evidence for the disposition effect, i.e. home owners do not sell their houses even if they get the signal of declining house prices from neighboring regions. Thus, our results confirm the findings by Genesove and Mayer (2001) and Engelhardt (2003) of loss aversion. However, the authors find only nominal and not real loss aversion as we do. As mentioned in the introduction, homeowners tend to avoid new information regarding neighboring house price declines or find explanations as to why neighboring house price decreases will not spill over into their region. This behavior indeed leads to smaller spillover effects in times of decreasing neighboring house prices.

Overall, we find evidence for heterogeneity in spatial spillovers of house price developments across space and time. This confirms and augments the findings by Gray (2012) of heterogeneity in house price spillovers across space.

Figure 3.4 is a plot of the individual coefficients of the spatially lagged depen-

dent variable for four points in time and reveals the amount of heterogeneity across time and space in spatial house price spillovers. The plotted coefficients vary between $\rho_0 = 0.33$, implying weak spatial spillovers of house price developments, and $\rho_0 + \rho_1 = 0.60$, implying strong spatial spillovers of house price developments. Those strong spillovers are especially observed for the coastal regions for example in 2004Q2, or 2005Q3. The house price spillovers in the inland at that time are more at the lower bound of 0.33. When house prices started to decline at the end of 2006, the spillovers quickly declined in the coastal regions and reached for almost all regions the lower bound of 0.33 in 2008Q1. Moreover, 4192 out of 6699 observations overall (319 regions over 21 time periods) have coefficients between 0.4 and 0.5, as compared to the extreme regime coefficients $\rho_0=0.33$ and $\rho_0 + \rho_1=0.6$. Thus, assuming only two extreme regimes without smooth transition between them would not be appropriate for more than 60 percent of all observations.

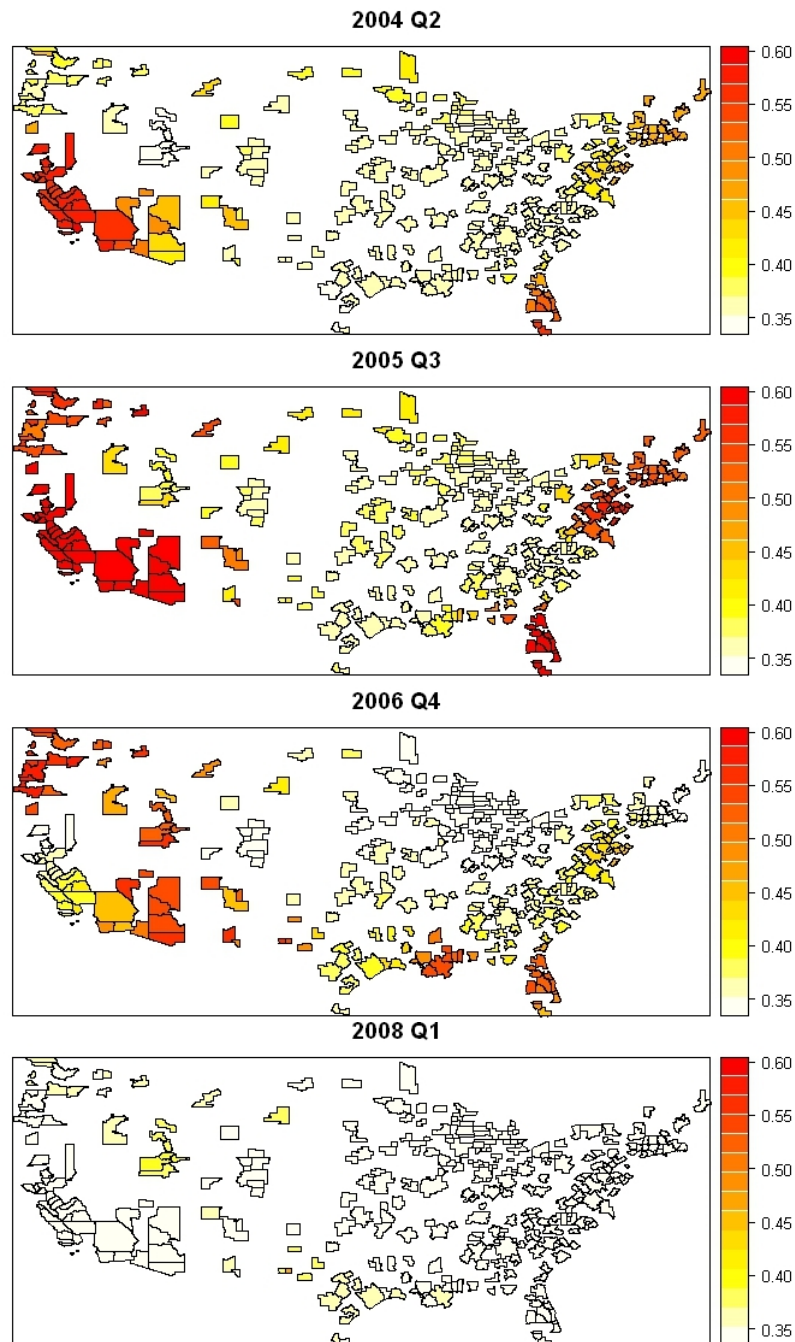
The estimation results further reveal that there is no heterogeneity in the effect of real per capita disposable income on house price developments, as the test statistic reveals no significant difference between the coefficients ζ_0 and $\zeta_0 + \zeta_1$. The coefficient on the unemployment rate is significant negative in one extreme regime and significant positive in the other extreme regime. The coefficients between β_0 and $\beta_0 + \beta_1$ are close to zero and probably insignificant, as in the spatial panel regressions. Therefore we do not worry about the unreasonable positive effect of the unemployment rate in times of very high house price growth rates. Moreover, the coefficient of population growth is insignificant in times of decreasing neighboring house prices and becomes significant when neighboring house prices increase. This result implies that, compared to the strong increases, the strong decline in house prices in the sample period could less be explained by fundamentals. In sum, heterogeneity in the effect of the fundamentals on house price dynamics could not be detected for all variables; real per capita disposable income and the unemployment rate have a homogeneous effect across time and space.

Table 3.3: PSTR estimation results, two extreme regimes

Dependent variable: quarterly real house price growth rate, hpi			
Variable	Coefficients	Huber-White t-statistic	Difference between coefficients
α_0 population $_{t-3}$	0.07	1.43	
β_0 unemployment $_{t-2}$	-0.02***	-11.34	
ζ_0 income $_{t-2}$	0.13***	7.46	
δ_0 building permits $_{t-1}$	0.005***	9.41	
ρ_0 W \times hpi $_{t-1}$	0.33***	9.39	
$\alpha_0 + \alpha_1$ population $_{t-3}$	0.50***	8.60	-5.73***
$\beta_0 + \beta_1$ unemployment $_{t-2}$	0.05***	9.95	-11.41***
$\zeta_0 + \zeta_1$ income $_{t-2}$	0.18***	5.36	-1.40
$\delta_0 + \delta_1$ building permits $_{t-1}$	0.002***	3.76	3.97***
$\rho_0 + \rho_1$ W \times hpi $_{t-1}$	0.60***	17.37	-5.44***
γ , slope	37.49		
c , location parameter	0.07		
R^2	0.63		
AIC	-8.3949		
Observations	6699		
Number of groups	319		
Time period	2004Q2-2009Q2		
Time dummies	yes		

Notes: Fixed effects estimation with robust standard errors to conditional heteroskedasticity of unknown form; ***, **, * statistically significant at one, five, and ten percent, respectively. Variables: *population* = annual population growth, *unemployment* = annual growth of the unemployment rate, *income* = annual real per capita disposable income growth, *building permits* = log building permits per population, *W \times hpi* = spatially weighted quarterly real house price growth rate of neighboring regions.

Figure 3.4: Individual spatial spillover parameters at different points in time, two extreme regimes



Notes: Individual spatial spillover parameters vary between $\rho_0 = 0.33$ and $\rho_0 + \rho_1 = 0.60$.

The test of no remaining non-linearity proposed by González et al. (2005) confirms that the linear model presented in Table 3.2 is not appropriate. Furthermore, the R^2 and the Akaike information criterion (AIC) confirm an improvement of the non-linear model over the linear model. However, test results reveal that there is still non-linearity in the non-linear model with two extreme regimes and also in the model with three extreme regimes. González et al. (2005) argue that heteroskedasticity may lead to a higher test statistic and thus to reject the null hypothesis of no remaining non-linearity more often. We conclude that, even if there is room for improvement, the PSTR model with two extreme regimes is a first step to model non-linearities, and thus, to better understand regional house price dynamics.

3.6 Conclusion

This paper is a joint analysis of three spatial characteristics in house price dynamics, namely spatial dependence, spatial heterogeneity, and heterogeneity in spatial dependence. While spatial dependence and spatial heterogeneity are well established aspects of house price developments, heterogeneity in spatial dependence has not gained much attention yet. We argue that the disposition effect may explain different house price spillovers across space and time. Assuming incorrectly homogeneous spillovers could locally give a misleading picture of house price dynamics.

First, a spatial panel regression is estimated to see whether there is overall spatial dependence in house price developments. Subsequently, a panel smooth transition regression model is applied to estimate the heterogeneity across space and time in spatial dependence and in the effect of the fundamentals on house price dynamics. To the best of our knowledge, this paper is the first applying this nonlinear model to jointly analyze the three aspects of spatial data, the first paper which explicitly models heterogeneity in spatial dependence, and the first paper which tries to model the disposition effect using heterogeneity in spatial spillovers.

The results reveal strong house price spillovers when the average annual house price increase of the neighboring regions is greater than 15 percent. Significant lower house price spillovers are detected for times of declining house prices in the neighboring regions. This is seen as evidence for the disposition effect, i.e. that people hold on their losing assets even if they get strong signals of declining house prices from the neighboring regions. Thus our results confirm previous findings by Genesove and Mayer (2001) and Engelhardt (2003) of loss aversion in the housing market. Heterogeneity in the effect of the fundamentals on house price dynamics is only found for population growth and building permits, but not for real per capita disposable income and the unemployment rate. The detected heterogeneity in the effect of population growth on house price developments suggest that fundamentals serve less explaining the house price developments in times of declining house prices compared to strongly increasing house prices.

This analysis shows that it is not appropriate to assume uniform house price spillovers across space and time. In times of declining house prices the spillovers are much lower than what the linear estimation suggests. The panel smooth transition regression model is an appropriate tool to model those nonlinearities in spatial spillovers across time and space.

3.A Appendix

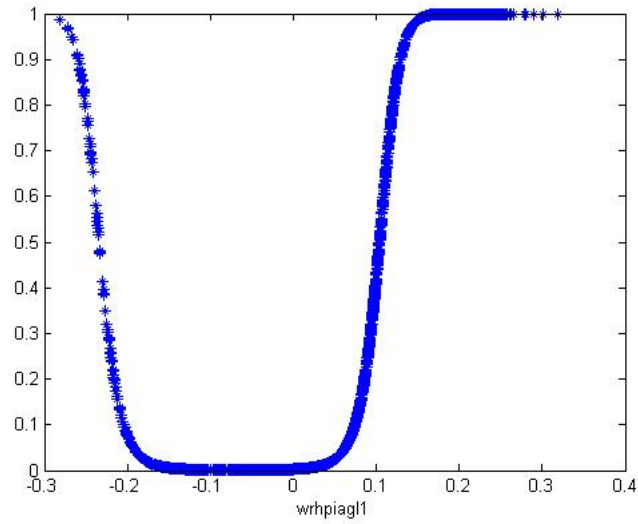
Because of theoretical consideration and the fact that the estimation is conducted over a relative short time span of 21 quarters, we assumed that there are only two extreme regimes. However, Figure 3.2, which is a plot of our transition variable, the spatially weighted annual growth rate of neighboring house prices, $Whpiag_{t-1}$, shows that in the sample there are very high annual growth rates of above 20 percent and very low growth rates of below minus 20 percent. This could imply that there are three regimes, one for very high growth rates of the transition variable, one for strong negative growth rates of the transition variable, and one for growth rates in between. That is why we also estimate the PSTR model for the case of three extreme regimes. For the case of three extreme regimes, the logistic transition function is given by:

$$g(q_{it}; \gamma, c) = \frac{1}{1 + e^{-\gamma(q_{it}-c_1)(q_{it}-c_2)}}, \quad (3.10)$$

where $c = (c_1, c_2)$ is the vector of location parameters. If the transition variable, q_{it} , is smaller than the first location parameter, c_1 , the transition function, $g(q_{it}; \gamma, c)$, tends to one, and the coefficients tend to $\theta_0 + \theta_1$. If the transition variable, q_{it} , is larger than the location parameter, c_2 , the transition function, $g(q_{it}; \gamma, c)$, tends again to one, and the coefficients tend to $\theta_0 + \theta_1$. If the transition variable, q_{it} , is larger than the first location parameter, c_1 , and smaller than the second location parameter, c_2 , the transition function, $g(q_{it}; \gamma, c)$, tends to zero, and the coefficients tend to θ_0 .

The corresponding transition function for the case of three extreme regimes is plotted in Figure 3.5. It appears that the assumption of three extreme regimes is not appropriate. The transition function is similar to the transition function in case of two extreme regimes between 20 percent decrease and 30 percent increase in neighboring house prices. Between 30 and 20 percent decrease there is a smooth transition of the transition function from 1 to 0, however, there are no observations in the extreme regime of strongly declining house prices. Thus, it appears to be appropriate to stick to the estimation results presented in Table 3.3. However, the results for three regimes are similar to the results presented in Table 3.3 and will be provided upon request.

Figure 3.5: Transition function vs. transition variable, three extreme regimes



Notes: Transition variable - spatially weighted annual real house price growth rate of neighboring regions, $Whpiag_{t-1}$

Bibliography

- Acs, Z. and C. Armington (2004). Employment Growth and Entrepreneurial Activity in Cities. *Regional Studies* 38(8), 911–927.
- Acs, Z., D. Audretsch, and D. Evans (1994). Why Does the Self-Employment Rate Vary Across Countries and Over Time? CEPR Discussion Paper 871, Centre for Economic Policy Research.
- Acs, Z., P. Braunerhjelm, D. Audretsch, and B. Carlsson (2009). The Knowledge Spillover Theory of Entrepreneurship. *Small Business Economics* 32(1), 15–30.
- Acs, Z., S. Desai, and J. Hessels (2008). Entrepreneurship, economic development and institutions. *Small Business Economics* 31(3), 219–234.
- Acs, Z. and L. Szerb (2009). The Global Entrepreneurship Index (GEINDEX). *Foundations and Trends in Entrepreneurship* 5(5), 341–435.
- Acs, Z. and A. Varga (2005). Entrepreneurship, Agglomeration and Technological Change. *Small Business Economics* 24(3), pp. 323–334.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- Anselin, L. (2002). Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models. *Agricultural Economics* 27(3), 247–267.
- Anselin, L., J. Gallo, and H. Jayet (2008). Spatial Panel Econometrics. In L. Mátyás and P. Sevestre (Eds.), *The Econometrics of Panel Data*, Volume 46 of *Advanced Studies in Theoretical and Applied Econometrics*, pp. 625–660. Springer Berlin Heidelberg.
- Anselin, L., A. Varga, and Z. Acs (1997). Local Geographic Spillovers between University Research and High Technology Innovations. *Journal of Urban Economics* 42(3), 422–448.

- Armington, C. and Z. Acs (2002). The Determinants of Regional Variation in New Firm Formation. *Regional Studies* 36(1), 33–45.
- Arrow, K. J. (1962). Economic Welfare and the Allocation of Resources for Invention. In R. R. Nelson (Ed.), *The rate and direction of inventive activity*, pp. 609–626. Princeton University Press.
- Audretsch, D. and M. Feldman (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review* 86(3), 630–640.
- Audretsch, D. and M. Fritsch (2002). Growth Regimes over Time and Space. *Regional Studies* 36(2), 113–124.
- Audretsch, D. and M. Keilbach (2004). Entrepreneurship Capital and Economic Performance. *Regional Studies* 38(8), 949–959.
- Audretsch, D. and E. E. Lehmann (2005). Does the Knowledge Spillover Theory of Entrepreneurship hold for Regions? *Research Policy* 34(8), 1191–1202.
- Autant-Bernard, C., P. Billand, D. Frachisse, and N. Massard (2007). Social Distance versus Spatial Distance in R&D Cooperation: Empirical Evidence from European Collaboration Choices in Micro and Nanotechnologies. *Papers in Regional Science* 86(3), 495–519.
- Badinger, H. and T. Url (2002). Determinants of Regional Unemployment: Some Evidence from Austria. *Regional Studies* 36(9), 977–988.
- Baldwin, J., D. Beckstead, W. M. Brown, and D. Rigby (2008). Agglomeration and the Geography of Localization Economies in Canada. *Regional Studies* 42(1), 117–132.
- Barro, R. and J. Lee (1993). International Comparisons of Educational Attainment. *Journal of Monetary Economics* 32(3), 363–394.
- Barro, R. and X. Sala-i-Martin (1995). *Economic Growth*. McGraw-Hill.

- Beck, N., K. Gleditsch, and K. Beardsley (2006). Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy. *International Studies Quarterly* 50(1), 27–44.
- Berkowitz, D. and D. N. DeJong (2005). Entrepreneurship and Post-socialist Growth. *Oxford Bulletin of Economics and Statistics* 67(1), 25–46.
- Beugelsdijk, S. (2007). Entrepreneurial culture, regional innovativeness and economic growth. *Journal of Evolutionary Economics* 17(2), 187–210.
- Bhattacharjee, A. and C. Jensen-Butler (2006). Estimation of Spatial Weights Matrix, with an Application to Diffusion in Housing Demand. CRIEFF Discussion Papers 519, Centre for Research into Industry, Enterprise, Finance and the Firm.
- Bivand, R. and D. Yu (2011). *Package 'spgwr' Geographically weighted regression (Version 0.6-13)*.
- Blanchflower, D. G. (2000). Self-employment in OECD countries. *Labour Economics* 7(5), 471–505.
- Bosma, N., G. Wit, and M. Carree (2005). Modelling Entrepreneurship: Unifying The Equilibrium and Entry/Exit Approach. *Small Business Economics* 25(1), 35–48.
- Bottazzi, L. and G. Peri (2003). Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review* 47(4), 687–710.
- Brady, R. R. (2011). Measuring the diffusion of housing prices across space and over time. *Journal of Applied Econometrics* 26(2), 213–231.
- Braunerhjelm, P., Z. Acs, D. Audretsch, and B. Carlsson (2010). The missing link: knowledge diffusion and entrepreneurship in endogenous growth. *Small Business Economics* 34(2), 105–125.
- Breinlich, H. (2006). The spatial income structure in the European Union - what role for Economic Geography? *Journal of Economic Geography* 6(5), 593–617.

- Breitenecker, R. J. and R. Harms (2010). Dealing With Spatial Heterogeneity in Entrepreneurship Research. *Organizational Research Methods* 13(1), 176–191.
- Brunsdon, C., A. S. Fotheringham, and M. E. Charlton (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis* 28(4), 281–298.
- Caliendo, M. and A. S. Kritikos (2010). Start-ups by the unemployed: characteristics, survival and direct employment effects. *Small Business Economics* 35(1), 71–92.
- Capello, R. and P. Nijkamp (2009). Introduction: Regional Growth and Development Theories in the Twenty-first Century - Recent theoretical Advances and Future Challenges. In R. Capello and P. Nijkamp (Eds.), *Handbook of Regional Growth and Development Theories*, pp. 1–18. Edward Elgar Publishing Limited.
- Carree, M., A. van Stel, R. Thurik, and S. Wennekers (2002). Economic Development and Business Ownership: An Analysis Using Data of 23 OECD Countries in the Period 1976-1996. *Small Business Economics* 19(2), 271–290.
- Carree, M., A. van Stel, R. Thurik, and S. Wennekers (2007). The relationship between economic development and business ownership revisited. *Entrepreneurship & Regional Development: An International Journal* 19(3), 281–291.
- Case, A. C., H. S. Rosen, and J. R. Hines (1993). Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States. *Journal of Public Economics* 52(3), 285–307.
- Chew, R. and L. Tan (1999). *Human Capital Formation as an Engine of Growth: The East Asian Experience*. Institute of Southeast Asian Studies.
- de Bandt, O., K. Barhoumi, and C. Bruneau (2010). The International Transmission of House Price Shocks. In O. de Bandt, T. Knetsch, J. Penalosa,

- and F. Zollino (Eds.), *Housing Markets in Europe*, pp. 129–158. Springer Berlin Heidelberg.
- de Bandt, O. and S. Malik (2010). Is there Evidence of Shift-Contagion in International Housing Markets? Banque de France Working Paper 295, Banque de France.
- Dejardin, M. and M. Carree (2011). Firm Entry and Exit in Local Markets: Market Pull and Unemployment Push. Working Papers 1114, University of Namur, Department of Economics.
- Delgado, M., M. E. Porter, and S. Stern (2010). Clusters and Entrepreneurship. *Journal of Economic Geography* 10(4), 495–518.
- Dieleman, F. M., W. A. V. Clark, and M. C. Deurloo (2000). The Geography of Residential Turnover in Twenty-seven Large US Metropolitan Housing Markets, 1985-95. *Urban Studies* 37(2), 223–245.
- Doering, T. and J. Schnellenbach (2006). What do we know about Geographical Knowledge Spillovers and Regional Growth?: A Survey of the Literature. *Regional Studies* 40(3), 375–395.
- Driscoll, J. C. and A. C. Kraay (1998). Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data. *The Review of Economics and Statistics* 80(4), 549–560.
- Engelhardt, G. V. (2003). Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States. *Journal of Urban Economics* 53(1), 171–195.
- European Commission and Eurostat (1999). *Handbook on quarterly national accounts, Theme 2, Economy and finance; Methods and nomenclatures*. Luxembourg : Office for Official Publications of the European Communities.
- Faggio, G. and O. Silva (2012). Does Self-Employment Measure Entrepreneurship? Evidence from Great Britain. SERC Discussion Papers 0109, Spatial Economics Research Centre, LSE.

- Federal Register (2010, June). 2010 Standards for Delineating Metropolitan and Micropolitan Statistical Areas. *Federal Register* 75(123), 37246–37252.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.
- Fischer, M. M., M. Bartkowska, A. Riedl, S. Sardadvar, and A. Kunnert (2009). The Impact of Human Capital on Regional Labor Productivity in Europe. *Letters in Spatial and Resource Sciences* 2(2), 97–108.
- Fischer, M. M. and P. Nijkamp (2009). Entrepreneurship and Regional Development. In R. Capello and P. Nijkamp (Eds.), *Handbook of Regional Growth and Development Theories*, pp. 182–200. Edward Elgar Publishing Limited.
- Fölster, S. (2000). Do Entrepreneurs Create Jobs? *Small Business Economics* 14(2), 137–148.
- Fornahl, D. (2003). Entrepreneurial Activities in a Regional Context. In D. Fornahl and T. Brenner (Eds.), *Cooperation, Networks, and Institutions in Regional Innovations Systems*, pp. 38–57. Edward Elgar Publishing Limited.
- Fotheringham, A. S. (2009). Geographically Weighted Regression. In A. S. Fotheringham and P. Rogerson (Eds.), *The Sage Handbook of Spatial Analysis*. SAGE Publications Limited.
- Fotheringham, A. S., C. Brunson, and M. E. Charlton (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Wiley.
- Fritsch, M. (2008). How does new business formation affect regional development? Introduction to the special issue. *Small Business Economics* 30(1), 1–14.
- Fritsch, M. and P. Müller (2004). Effects of New Business Formation on Regional Development over Time. *Regional Studies* 38(8), 961–975.
- Fritsch, M. and A. Schroeter (2011). Why does the effect of new business formation differ across regions? *Small Business Economics* 36(4), 383–400.

- Fritsch, M. and M. Wyrwich (2012). The Long Persistence of Regional Entrepreneurship Culture - Germany 1925-2005. DIW Discussion Papers 1224, DIW Berlin.
- Funke, M. and A. Niebuhr (2005). Spatial R&D Spillovers and Economic Growth - Evidence from West Germany. *Regional Studies* 39(1), 143–153.
- Genesove, D. and C. Mayer (2001). Loss Aversion and Seller Behavior: Evidence from the Housing Market. *The Quarterly Journal of Economics* 116(4), 1233–1260.
- Giussani, B. and G. Hadjimatheou (1991). Modeling Regional House Prices in the United Kingdom. *Papers in Regional Science* 70(2), 201–219.
- González, A., T. Teräsvirta, and D. van Dijk (2005). Panel Smooth Transition Regression Models. Working Paper Series in Economics and Finance 604, Stockholm School of Economics.
- Gray, D. (2012). District House Price Movements in England and Wales 1997-2007: An Exploratory Spatial Data Analysis Approach. *Urban Studies* 49(7), 1411–1434.
- Harris, R., J. Moffat, and V. Kravtsova (2011). In Search of 'W'. *Spatial Economic Analysis* 6(3), 249–270.
- Holcombe, R. (2006). *Entrepreneurship And Economic Progress*. Foundations of the Market Economy Series. Routledge.
- Holly, S., M. H. Pesaran, and T. Yamagata (2011). The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics* 69(1), 2–23.
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 221–233.
- Jaffe, A. B. (1986). Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value. *The American Economic Review* 76(5), 984–1001.

- Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2), 263–291.
- Kodres, L. E. and M. Pritsker (2002). A Rational Expectations Model of Financial Contagion. *The Journal of Finance* 57(2), 769–799.
- Kosfeld, R. (2007). Regional Spillovers and Spatial Heterogeneity in Matching Workers and Employers in Germany. *Jahrbücher für Nationalökonomie und Statistik* 227(3), 236–253.
- Kuethé, T. and V. Pede (2011). Regional Housing Price Cycles: A Spatio-temporal Analysis Using US State-level Data. *Regional Studies* 45(5), 563–574.
- Lee, C.-C. and M.-S. Chien (2011). Empirical Modelling of Regional House Prices and the Ripple Effect. *Urban Studies* 48(10), 2029–2047.
- Lee, L.-f. and J. Yu (2010). Some recent developments in spatial panel data models. *Regional Science and Urban Economics* 40(5), 255–271.
- LeSage, J. P. and M. M. Fischer (2008). Spatial Growth Regressions: Model Specification, Estimation and Interpretation. *Spatial Economic Analysis* 3(3), 275–304.
- LeSage, J. P. and R. K. Pace (2009). *Introduction to Spatial Econometrics*. Chapman & Hall/CRC.
- LeSage, J. P. and R. K. Pace (2010). Spatial Econometric Models. In M. M. Fischer and A. Getis (Eds.), *Handbook of Applied Spatial Analysis*, pp. 355–376. Springer Berlin Heidelberg.
- LeSage, J. P. and W. Polasek (2008). Incorporating Transportation Network Structure in Spatial Econometric Models of Commodity Flows. *Spatial Economic Analysis* 3(2), 225–245.
- Leung, Y., C.-L. Mei, and W.-X. Zhang (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A* 32(1), 9–32.

- Li, H., S. Cheng, and K. E. Haynes (2011). The Employment Effects of New Business Formation: A Regional Perspective. *Economic Development Quarterly* 25(3), 282–292.
- Lucas Jr., R. E. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics* 9(2), 508–523.
- Lucas Jr., R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics* 22(1), 3–42.
- Mankiw, N. G., D. Romer, and D. N. Weil (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics* 107(2), 407–437.
- Meen, G. (1999). Regional House Prices and the Ripple Effect: A New Interpretation. *Housing Studies* 14(6), 733–753.
- Mikhed, V. and P. Zemcik (2009). Do house prices reflect fundamentals? Aggregate and panel data evidence. *Journal of Housing Economics* 18(2), 140–149.
- Molloy, R., C. L. Smith, and A. K. Wozniak (2011). Internal Migration in the United States. NBER Working Paper 17307, National Bureau of Economic Research.
- Müller, P. (2006). Exploring the Knowledge Filter: How Entrepreneurship and University– industry Relationships drive Economic Growth. *Research Policy* 35(10), 1499–1508.
- Müller, S. (2012). Identifying spatial nonstationarity in German regional firm start-up data. *Jahrbuch für Regionalwissenschaft* 32(2), 113–132.
- Niebuhr, A. and S. Stiller (2004). Integration Effects in Border Regions - A Survey of Economic Theory and Empirical Studies. *Jahrbuch für Regionalwissenschaft* 24(1), 3–21.
- Nistor, A. (2009). Assessing the Effectiveness of Human Capital Investments on the Regional Unemployment Rate in the United States: 1990 and 2000. *International Regional Science Review* 32(1), 65–91.

- OECD (2000). The Partial Renaissance of the Self-Employed. In *OECD Employment Outlook*, Chapter 5, pp. 155–199. Paris: OECD.
- Olejnik, A. (2008). Using the Spatial Autoregressively Distributed Lag Model in Assessing the Regional Convergence of Per-Capita Income in the EU25. *Papers in Regional Science* 87(3), 371–384.
- Parent, O. and J. P. LeSage (2008). Using the Variance Structure of the Conditional Autoregressive Spatial Specification to Model Knowledge Spillovers. *Journal of Applied Econometrics* 23(2), 235–256.
- Parent, O. and J. P. LeSage (2010). A spatial dynamic panel model with random effects applied to commuting times. *Transportation Research Part B: Methodological* 44(5), 633–645.
- Partridge, M. D., D. S. Rickman, K. Ali, and M. R. Olfert (2008). The Geographic Diversity of U.S. Nonmetropolitan Growth Dynamics: A Geographically Weighted Regression Approach. *Land Economics* 84(2), 241–266.
- Patuelli, R., A. Reggiani, P. Nijkamp, and F.-J. Bade (2009). Spatial and Commuting Networks: A Unifying Perspective. In A. Reggiani and P. Nijkamp (Eds.), *Complexity and Spatial Networks: In Search of Simplicity*, pp. 257–271. Springer Berlin Heidelberg.
- Pfeiffer, F. and F. Reize (2000). Business start-ups by the unemployed - an econometric analysis based on firm data. *Labour Economics* 7(5), 629–663.
- Porter, M. E. (1998a). Clusters and the New Economics of Competition. *Harvard Business Review* 76(6), 77–90.
- Porter, M. E. (1998b). *On Competition*. Harvard Business School.
- Porter, M. E. (2003). The Economic Performance of Regions. *Regional Studies* 37(6–7), 545–546.
- Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography* 2(4), 373–406.

- Reggiani, A., P. Bucci, and G. Russo (2011). Accessibility and Network Structures in the German Commuting. *Networks and Spatial Economics* 11(4), 621–641.
- Rocha, H. O. and R. Sternberg (2005). Entrepreneurship: The Role of Cluster Theoretical Perspectives and Empirical Evidence from Germany. *Small Business Economics* 24(3), 267–292.
- Rodríguez-Pose, A. and R. Crescenzi (2008). Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. *Regional Studies* 42(1), 51–67.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy* 98(S5), 71–102.
- Sanandaji, T. (2010). Self-Employment Does Not Measure Entrepreneurship. Job market paper, University of Chicago, The Institute For Industrial Economics.
- Scherngell, T. and M. J. Barber (2009). Spatial Interaction Modelling of Cross-Region R&D Collaborations: Empirical Evidence from the 5th EU Framework Programme. *Papers in Regional Science* 88(3), 531–546.
- Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*. Transaction Publishers.
- Shearmur, R., P. Apparicio, P. Lizion, and M. Polèse (2007). Space, Time, and Local Employment Growth: An Application of Spatial Regression Analysis. *Growth and Change* 38(4), 696–722.
- Shefrin, H. and M. Statman (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance* 40(3), 777–790.
- Shiller, R. J. (1999). Human behavior and the efficiency of the financial system. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1, Chapter 20, pp. 1305–1340. Elsevier.

- Sorenson, O. and P. G. Audia (2000). The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940-1989. *American Journal of Sociology* 106(2), 424–462.
- Sutter, R. (2010). *The Psychology of Entrepreneurship and the Technological Frontier—A Spatial Econometric Analysis of Regional Entrepreneurship in the United States*. Ph. D. thesis, George Mason University.
- Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association* 8(4), 677–716.
- Teräsvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. *Journal of the American Statistical Association* 89(425), 208–218.
- Thaler, R. H. (1999). Mental Accounting Matters. *Journal of Behavioral Decision Making* 12(3), 183–206.
- Thornton, P. H., D. Ribeiro-Soriano, and D. Urbano (2011). Socio-cultural factors and entrepreneurial activity. *International Small Business Journal* 29(2), 105–118.
- Thurik, R., M. Carree, A. van Stel, and D. Audretsch (2008). Does self-employment reduce unemployment? *Journal of Business Venturing* 23(6), 673–686.
- Trupkin, D. and R. Ibarra (2011). The Relationship between Inflation and Growth: A Panel Smooth Transition Regression Approach for Developed and Developing Countries. Documentos de Trabajo/Working Papers 1107, Facultad de Ciencias Empresariales y Economía. Universidad de Montevideo.
- van Dijk, B., P. H. Franses, R. Paap, and D. van Dijk (2011). Modelling regional house prices. *Applied Economics* 43(17), 2097–2110.

- van Stel, A., M. Carree, and R. Thurik (2005). The Effect of Entrepreneurial Activity on National Economic Growth. *Small Business Economics* 24(3), 311–321.
- van Stel, A. and D. Storey (2004). The Link between Firm Births and Job Creation: Is there a Upas Tree Effect? *Regional Studies* 38(8), 893–909.
- Verheul, I., S. Wennekers, D. Audretsch, and R. Thurik (2002). An Eclectic Theory of Entrepreneurship: Policies, Institutions and Culture. In D. Audretsch, R. Thurik, I. Verheul, and S. Wennekers (Eds.), *Entrepreneurship: Determinants and Policy in a European-US Comparison*, Volume 27 of *Economics of Science, Technology and Innovation*, Chapter 2, pp. 11–81. Kluwer Academic Publishers.
- Wennekers, S., A. van Stel, M. Carree, and R. Thurik (2010). The relationship between entrepreneurship and economic development: is it U-shaped? *Foundations and Trends in Entrepreneurship* 6(3), 167–237.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48(4), 817–838.
- Wood, R. (2003). The Information Content of Regional House Prices: Can they be used to Improve National House Price Forecasts? *Bank of England Quarterly Bulletin* Autumn, 304–314.
- Yu, D. (2011). Exploring Spatiotemporally Varying Regressed Relationships: The Geographically Weighted Panel Regression Analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38, 134–139.

List of Publications and Working Papers

Publications and Working Papers part of the Dissertation

Chapter	Publication or Working Paper Title	Coauthors	Authors' Contribution
I	Do Regions with Entrepreneurial Neighbours Perform Better? A Spatial Econometric Approach for German Regions, <i>Regional Studies</i> , forthcoming. http://dx.doi.org/10.1080/00343404.2012.697143	Konstantin A. Kholodilin	Main responsibility for research design, literature review, data analysis, statistical inference, interpretation of the results, and writing of the paper. The coauthor constructed 33 of the 56 weight matrices used.
II	Self-Employment and Economic Performance -A Geographically Weighted Regression Approach for European Regions, DIW Discussion Papers 1272 (2013).	none	Author's independent research
III	The Spatial Dimension of US House Price Developments, DIW Discussion Papers 1270 (2013).	none	Author's independent research

Publications and Working Papers not part of the Dissertation

Publication or Working Paper Title	Coauthors
The Global Dimension of Inflation - Evidence from Factor-Augmented Phillips Curves, <i>Oxford Bulletin of Economics and Statistics</i> , 75 (2013), 103-122.	Sandra Eickmeier
Business Cycles, Unemployment and Entrepreneurial Entry - First Evidence from Germany, DIW Discussion Papers 1281 (2013).	Michael Fritsch, Alexander S. Kritikos

Curriculum Vitae

The CV is not included in the online version for reasons of data protection.

Curriculum Vitae

The CV is not included in the online version for reasons of data protection.

Curriculum Vitae

The CV is not included in the online version for reasons of data protection.

Curriculum Vitae

The CV is not included in the online version for reasons of data protection.

Curriculum Vitae

The CV is not included in the online version for reasons of data protection.

Curriculum Vitae

The CV is not included in the online version for reasons of data protection.

Ehrenwörtliche Erklärung

Ich habe die vorgelegte Dissertation selbst verfasst und dabei nur die von mir angegebenen Quellen und Hilfsmittel benutzt. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, sowie alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Katharina Pijnenburg