Regional Diversity and Entrepreneurship in Germany

by

David Audretsch, Dirk Dohse and Annekatrin Niebuhr

Preliminary, not for citation

Paper to be presented at the international workshop
Agglomeration and Growth in Knowledge-based Societies

Kiel, Germany, April 20-21

1. Introduction

The start-up and running of new business is central to modern economies' dynamics and their ability to innovate and grow. Entrepreneurs, a species long-time neglected from mainstream economics, recently seem to regain the profession's attention, but they are no longer viewed as "lone giants" but rather as very interactive people who heavily depend on other people, resources and opportunities in their respective context.

If it is true that the spillover of new knowledge, which is arguably the most important input into the entrepreneurial process, is geographically localized then the regional and urban context should be particularly important determinants of entrepreneurship. Indeed, recent studies have shown the value of regional factors (such as agglomeration economies, regional R&D or regional income growth) in explaining differences in the entry rates of firms (Rosenthal and Strange 2003, Stuart and Sorenson 2003, Lee et al. 2004). It is, however, by no means clear whether and under which circumstances regional specialization or diversity is more conducive to entrepreneurship.

The current paper analyses the regional determinants of entrepreneurship as measured in terms of firm start-ups in Germany, Europe's largest economy. The German economy is characterized by an institutional background quite different from the US, a high share of foreign population (and immigrant entrepreneurs) and marked regional differences between (the former communist) East and West Germany. The purpose of our analysis is threefold:

- Firstly, we conjecture that the factors driving firm start-ups might differ considerably with respect to the *technology* (or knowledge) input necessary to start up a certain business. We therefore consider different kinds of start-ups (i.e. total start-ups, technology-oriented start-ups, start-ups in technology-oriented services and high tech start-ups) and investigate whether the factors that influence the variation in regional firm birth rates differ systematically with respect to the technology level of the respective start-ups.

- Secondly, the central focus of the paper is on investigating the role of *diversity* with respect to new firm foundation. Unlike earlier papers we take a broad approach and consider the significance of various dimensions of diversity for firm foundation, namely sectoral diversity, technological diversity, cultural diversity and occupational diversity. Our main hypothesis is that regional diversity is conducive to entrepreneurship, especially with respect to high tech start-ups.

- Finally, we analyse whether there are still systematic differences between the determinants of firm birth between the *Eastern and Western part* of the country more than 10 years after the start of the transformation process in East Germany.

The paper is organised as follows. Section 2 provides the theoretical background on firm formation and regional diversity and specifies the econometric model in its basic form. Section 3 contains a detailed description of the data set and a discussion of the variables used in the estimations. Section 4 presents and discusses the results of the econometric analysis. Section 5 concludes.

2. Diversity and Entrepreneurship: The Theoretical Background

Economists have long observed that entrepreneurial activity tends to vary systematically across geographic space (Carlton 1983, Storey 1991, Reynolds et al. 1994). In searching for a theoretical framework to provide a lens through which spatial variation of entrepreneurship could best be interpreted and explained, scholars have gravitated towards models highlighting the extent to which entrepreneurial opportunities prevail or are impeded within a spatial context. This has generated an exhaustive literature linking regional-specific characteristics reflecting characteristics alternatively promoting or impeding entrepreneurial opportunities to various measures of regional entrepreneurship. Most notably, regional-specific measures, such as growth, unemployment, population density, taxes, and industry structure have been found to influence the extent of entrepreneurial activity within a region.

More recently, the *knowledge spillover theory of entrepreneurship* (Acs et al., 2004, and Audretsch et al., 2006) was introduced to provide an explicit link between knowledge and entrepreneurship within the spatial context. The knowledge spillover theory of entrepreneurship posits that investments in knowledge by incumbent firms and research organizations such as universities will generate entrepreneurial opportunities because not all of the new knowledge will be pursued and commercialized by the incumbent firms. As Arrow

(1962) pointed out, new knowledge is inherently uncertain and asymmetric, so that incumbent firms and other organizations are unable to recognize and act upon all of the knowledge created by their own investments to generate that knowledge in the first place. What one (knowledge) worker perceives to be a potentially valuable idea may not actually be acknowledged as being valuable by the decision-making hierarchy of the firm. The knowledge filter (Acs et al., 2004 and Audretsch et al., 2006) refers to the extent that new knowledge remains uncommercialized by the organization creating that knowledge. It is the residual ideas and knowledge left uncommercialized that generate the opportunity for entrepreneurship. By pursuing ideas and knowledge created but left uncommercialized in an incumbent firm or organization to launch a new firm, the entrepreneurial venture serves as a conduit of knowledge spillovers. Compelling empirical evidence has found that new-firm start-ups are systematically greater in regions rich in knowledge than in regions poor in knowledge (Audretsch and Keilbach, 2007; Audretsch et al., 2006).[1] These studies implicitly assume that, given a certain investment in knowledge, economic agents will automatically identify and act upon entrepreneurial opportunities. That is, the capabilities of economic agents within the region to actually access and absorb the knowledge and ultimately utilize it to generate entrepreneurial activity was implicitly assumed to be invariant with respect to geographic space.

However, such an assumption violates one of the most significant insights by Jane Jacobs (1969, later echoed by Porter 1990) that regions with more diversity will facilitate the spillover of knowledge, which in turn should trigger more entrepreneurial activity. According to Jacobs, it is differences among people that foster looking at and appraising a given information set differently, thereby resulting in different appraisal of any new idea. After all, if all economic agents were homogeneous, or perfectly identical, a total consensus would reign with respect to any new idea, and there would be no reason to start a new firm. As Jacobs emphasized, it is differences across economic agents that lead to divergences in the valuation of new ideas, and it is these divergences in the value of ideas that trigger people to start a new venture.

Thus, while knowledge may be important to generate new ideas with an unknown distribution of unknown outcomes, it is the assessment of those new ideas by diverse economic agents characterized by differences in experiences, backgrounds, and capabilities

[1] Moreover, Audretsch and Dohse (2007) are able to show that being located in an agglomeration rich in knowledge resources is more conducive to firm growth than being located in a region that is less endowed with knowledge resources.

that leads to divergences in the valuation of such ideas which ultimately induce agents to resort to entrepreneurship to appropriate the value of their knowledge endowments.

This suggests that for knowledge spillovers to occur, more than investments in new knowledge is required. Rather, economic agents with the capabilities to access, absorb and commercialize that knowledge through the spillover conduit of entrepreneurship are also essential for generating knowledge spillovers. Diversity will enhance such entrepreneurial activity because diverse economic agents will value new ideas differently, leading them to respond to different ideas differently. It is this diversity in economic agents that triggers divergences in the evaluation of new ideas that is the basis for knowledge spillover entrepreneurship.

A recent literature has examined the spatial dimension of knowledge spillovers and found that they tend to be spatially localized within close geographic proximity to the knowledge source (Jaffe, 1989; Audretsch and Feldman, 1996; Audretsch and Stephan, 1996; Jaffe et al., 1993). Thus, those regions with more diversity would be expected to generate more entrepreneurial activity. By contrast, less diversity, or more homogeneity, would be expected to generate less entrepreneurship.

Glaeser et al. (1993) and Feldman and Audretsch (1999) argued and provided compelling evidence linking diversity to regional economic growth. However, in both studies, diversity was measured in terms of economic activity within the region, which reflects firm, but not in terms of the people actually inhabiting and working in the region. This misses the essential diversity argument by Jacobs, which is first and foremost about people and not necessarily firm. Thus, a major contribution of this paper is not only to link regional entrepreneurial activity to the extent of diversity characterizing that region, but also to measure diversity not just in terms of firms but also people.

Unlike earlier papers we take a broad approach and consider the significance of various dimensions of diversity for firm foundation, namely sectoral diversity, technological diversity, cultural diversity and occupational diversity. *Sectoral diversity* is probably the most common concept. The indicators of sectoral diversity used in this paper are calculated with employment shares of 28 industries documented in the appendix. We measure *technological diversity* by analysing the patent portfolio of each region, considering patent shares of 31 technological areas, also listed in the appendix. The indicator of technological diversity might reflect the heterogeneity of the regional knowledge stock. Moreover, two indicators that refer directly to the diversity in terms of the people are considered. We use information on regional employment by nationality (overall 213 nationalities) to calculate our measure of *cultural diversity*. The indicator accounts for both richness of the distribution (i.e. number of nationalities) and a relatively even distribution across nationalities. Thus, according to the measures, cultural diversity will increase if the number of nationalities rises or if the shares of different nationalities in employment converge. Our fourth indicator, measuring *occupational diversity* is based on regional employment data differentiated by occupation. The classification implies a differentiation into 83 occupations. Cultural and occupational diversity are supposed to capture diversity of economic agents that is expected to increase the exploitation of a given regional knowledge basis and thus promote entrepreneurial activity.

In order to arrive at robust results regarding the impact of different diversity measures the regression analysis departs from a model that includes a number of factors that have turned out to be important determinants of the regional firm birth rate in the empirical literature. These *control variables* include population density, industry density, the disposable income in the region under consideration and in neighbouring regions (spatially lagged exogenous variable), growth of disposable income, growth of disposable income in neighbouring regions, unemployment, share of foreign workers, percentage of highly qualified employees, R&D employment and indicators of the firm size structure of the region (shares of small and large firms in total employment).

The econometric model in its basic version has the form

$$(1)\quad SU_{it} = \alpha_0 + \sum_{l=1}^{2} \alpha_l KNOW_{lit} + \sum_{m=1}^{4} \beta_m DIV_{mit} + \sum_{n=1}^{N} \gamma_n CONTROL_{nit} + u_{it}$$

where SU_{it} is the start-up intensity (start-ups per 10.000 inhabitants) in region *i* and year *t*, DIV_{mit} is diversity measure *m* in region *i* and period *t*, $CONTROL_{nit}$ is control variable *n* in region *i* and period *t*. Moreover, we include two knowledge variables $KNOW_{lit}$, the share of R&D workers in total employment, and the percentage of highly skilled employees. The error term is denoted by u_{it} and assumed to be identically and independently distributed with mean μ_u and variance σ_u^2.

To check the robustness of results emerging from the pooled model given by equation (1), we apply additional regression models. Firstly, panel data models are used to control for unobserved time-invariant explanatory variables:

$$(2)\quad SU_{it} = \alpha_0 + \sum_{l=1}^{2} \alpha_l KNOW_{lit} + \sum_{m=1}^{4} \beta_m DIV_{mit} + \sum_{n=1}^{N} \gamma_n CONTROL_{nit} + \eta_i + \lambda_t + \nu_{it}$$

where η_i denotes a region-specific effect, controlling for unobservable regional characteristics that are time-invariant. λ_t captures unobservable time effects and ν_{it} is a white noise error term. We estimate fixed effects as well as random effects specifications.

Secondly, we allow for parameter heterogeneity by estimating spatial regimes models:

(3)

$$SU_{it} = \sum_{j=1}^{2} \left(\alpha_{0j} d_j + \sum_{l=1}^{2} \alpha_{lj} d_j KNOW_{lit} + \sum_{m=1}^{4} \beta_{mj} d_j DIV_{mit} + \sum_{n=1}^{N} \gamma_{nj} d_j CONTROL_{nit} + d_j \lambda_{jt} \right) + \eta_i + \nu_{it}$$

where d_j denotes a dummy variable for the subset *j* of the data. In this analysis, we differentiate two subsets (*j* = 1, 2), East and West German regions. A Wald test on equality of the coefficients is applied to investigate whether the differences between the slope estimates among East and West German regions are significant.

3. Data description

The cross section consists of 97 functional regions, so-called Raumordnungsregionen, which comprise several counties (NUTS 3 level) linked by intense commuting.[2] Thus, the observational units represent regional labour markets. Since this definition of regions does not account for other forms of economic activity such as consumption, we care for possible spillover effects caused e.g. by demand linkages and other kinds of spatial interaction.

[2] According to a definition by the German Federal Office for Building and Regional Planning (BBR) Raumordnungsregionen are intended to be comparable regions "that reflect in acceptable approximation the spatial and functional interrelation between core cities and their hinterland." (BBR 2001: 2).

Dependent variables

We measure regional entrepreneurship in terms of start-up intensity, i.e. start-ups per 10.000 inhabitants. As the annual variation in firm birth rates is high we follow the recommendation of the data provider (ZEW Mannheim) and use 4-year averages (1998-2001 and 2002-2005, respectively) of firm birth rates as dependent variables in the regression analyses. Unlike earlier papers we do not only consider total start ups but differentiate between firm birth rates at different technology levels. In the remainder of the paper we focus on four different groups of start-ups, namely total start-ups (*Su_all*), technology oriented start-ups (*Su_to*) which make out roughly 10 % of all start-ups, and two particularly interesting and important sub-groups of technology oriented start-ups, i.e. start-ups in technology oriented services (*Su_tos*) and high tech start-ups (*Su_ht*). A detailed description of the data set and the classification of start-ups (industrial sectors) according to their technology level can be found in the documentation by Metzger and Heger 2005.[3]

As the majority of start-ups are not technology oriented, the firm birth rate of low tech businesses (*Su_nto*) is highly correlated with the measure for overall firm foundation (*Su_all*). Moreover, the spatial pattern of firm birth rates at different technological levels is subject to a considerable variation, as shown by the correlation coefficients in Table 1. The correlation between total firm birth and technology oriented start-ups (*Su_to*) amounts to 0.66, whereas the coefficient between the overall rate of firm foundation and high tech start ups (*Su_ht*) is merely 0.13. There are also pronounced disparities within the class of technology oriented start-ups as shown by the correlation among new high tech firms and technology oriented services (0.42). This suggests that firm birth of different technological categories might be driven by different factors.

Table 1: Correlations of firm birth rates at different technology levels

	Su_all	Su_to	Su_ht	Su_tos	Su_nto
Su_all	1				
Su_to	0.65	1			
Su_ht	0.13	0.46	1		
Su_tos	0.66	0.99	0.42	1	
Su_nto	0.99	0.52	0.05	0.53	1

Source: ZEW Firm Foundation Report, own calculations.

[3] It should be noted that according to ZEW the start-up rates for some regions are upward biased due to regional differences in the data survey mode. We therefore exclude the affected regions (Hamburg, Braunschweig, Westpfalz and Rheinpfalz) from the database.

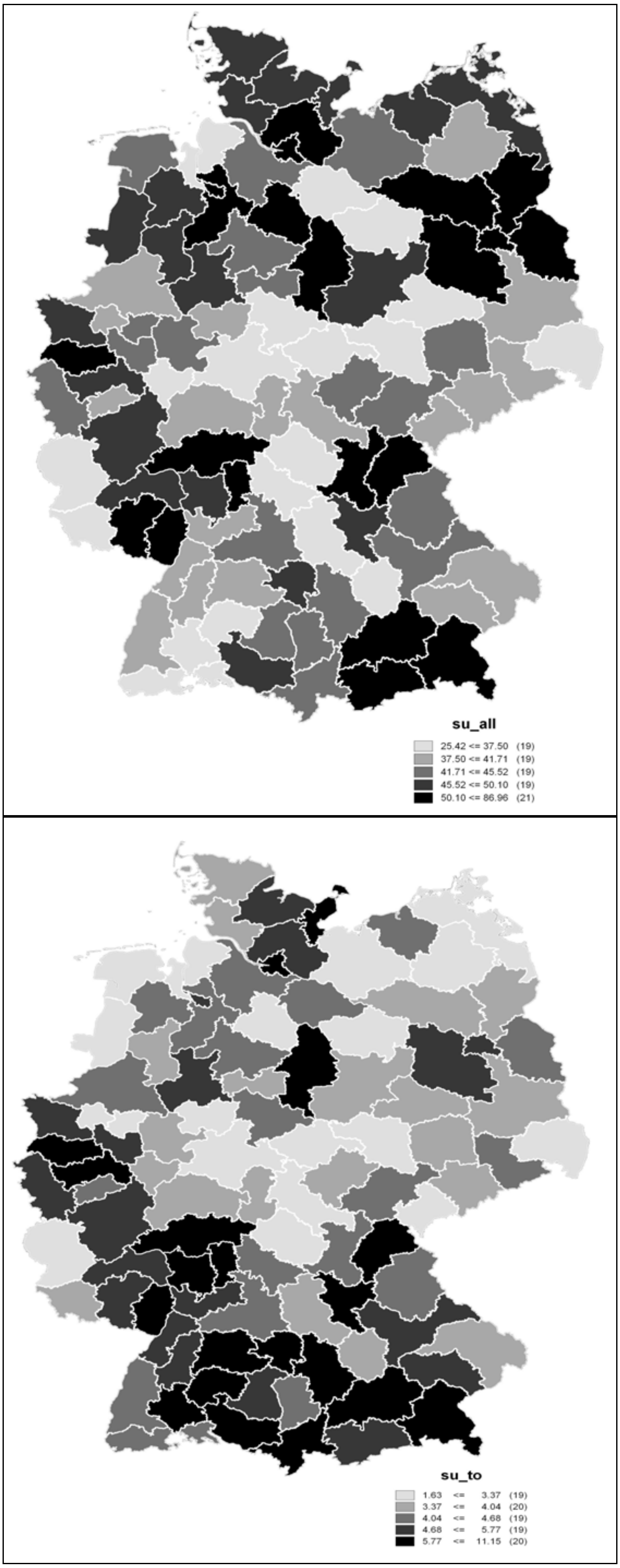
The substantial spatial variation of firm birth rates at different technological levels is also illustrated in map 1. While regions with a total start-up intensity (*Su_all*) in the highest 20 % can be found in the northern and eastern parts of the country as well as in the south and the west, we observe a striking concentration of regions with particularly high start-up intensities in technology-oriented industries (*Su_to*) in the southern parts of the country, i.e. in the states Bavaria, Baden-Württemberg and (the south of) Hesse.

Map A1 in the appendix shows that the spatial distribution of start-ups in technology oriented services (*Su_tos*) is quite similar to the spatial distribution of technology-oriented start-ups in general while the high tech start-up-rate (*Su_ht*) appears to be particularly high in the outmost south-west (Baden-Württemberg).

Explanatory Variables

In order to arrive at robust results regarding the impact of different diversity measures the regression analysis departs from a model that includes a number of factors that have turn out to be important determinants of the regional firm birth rate. We deal with potential endogeneity of some influential factors by using predetermined explanatory variables. Thus, as regards the average firm birth rate 1998-2001 (2002-2005) the explanatory variables refer to 1997 (2001) – unless stated otherwise.

Map 1: Regional distribution of total start-ups and technology-oriented start-ups



As mentioned above, the definition of spatial units is based on labour market criteria and does not consider other areas of economic activity. However, the size of the market for goods and services might well exceed the borders of the observational units. To capture effects linked to the size of the local market we include the disposable income of the regions (*INC*) and growth of disposable income (*INC_G*). Information on regional disposable income is available from the national accounts. Spillover effects resulting from a market size that does not correspond with our functional regions is taken into account by spatial lags of explanatory variables. The regression model is extended by spatial lags of disposable income (*W_INC*) and of the growth of disposable income (*W_INC_G*). Therefore, we investigate whether purchasing power in neighbouring regions has a significant impact on the firm birth rate. Modelling spillover effects requires some information on the structure of spatial interaction summarized by a spatial weight matrix. We apply two alternative specifications of weight matrices. The first specification involves a binary matrix such that the weights $w_{ij} = 1$ if the regions i and j share a border and $w_{ij} = 0$ otherwise. Secondly, a weighting scheme based on distance between region's capitals is considered. We fix a cut-off point of 100 kilometres, i.e. $w_{ij} = 0$ if the distance between i and j exceeds this threshold. The weights within the maximum range are calculated as the inverse of distance. All weight matrices are row standardized.

We further include the regional unemployment rate, i.e. number of unemployed as a percentage of the labor force (*UR*) as a major control variable. Conflicting hypotheses are discussed in the entrepreneurship literature regarding the impact of unemployment. Some authors argue that in case of high unemployment the propensity of people to start their own business might increase because of lacking alternative job opportunities. Based on this argument, we might expect that regions characterized by high unemployment rates realize high rates of new firm foundation. However, high unemployment can also indicate economic decline and low consumer demand. In this case a high rate of unemployment is likely to exert a dampening effect on entrepreneurship.

Moreover, we consider several explanatory variables in the regression model based on employment data provided by the German Federal Employment Agency. The employment statistic covers all employment subject to social security contributions.[4] The information refers to workplace location. We use employment data differentiated by nationality,

[4] Hence, civil servants and self-employed are not recorded in the employment statistic.

educational level, industry, occupation and firm size to generate several diversity measures and control variables that enter into the regression model.

The *knowledge spillover theory of entrepreneurship* stresses the importance of the regional knowledge basis for start-ups. Therefore, we consider two knowledge variables in the regression analysis. The share of highly qualified employees[5] in total employment (*HQ*) might exert a positive impact on regional start-ups for several reasons. Firstly, research on the individual characteristics of entrepreneurs indicates that people with high educational attainment are more likely to found a new firm than low skilled workers. Secondly, in particular high tech start-ups probably benefit from a large regional pool of highly skilled employees. This also applies to R&D employment, i.e. the share of R&D workers in total employment (*RD*).

Additional to these knowledge-based measures we include two variables that reflect the density of economic activity in general and are not directly knowledge (or human capital) oriented. The variable *PD* measures population density (inhabitants per square kilometer) in the German planning regions, whereas *ID* (=industry density) relates a region's manufacturing employment to the number of inhabitants.

We also include the share of foreign workers (*FP*) as an explanatory variable since some studies provide evidence on entrepreneurship to be affected by ethnic origin (e.g. Storey 1994). Empirical studies suggest that immigrants are more likely to be self-employed because they suffer from discrimination, lack contacts in existing businesses and are more willing to take risks such as starting a new business. Therefore regions marked by an above average share of foreign people are expected to show relatively high firm birth rates. However, this argument might only apply to total firm foundation and not to high tech start-ups.

To proxy the regional structure of establishment size we consider two variables that are based on regional employment data given for three categories of firm size. In the regression model the share of small firms (less than 20 employees; *SE*) in total employment and the percentage of employees in large firms (more than 500 employees; *LE*) are included. A high percentage

[5] Regionally disaggregated data on highly qualified employees are available from the German Federal Office for Building and Regional Planning (BBR). Highly qualified employees are – according to the definition used by the BBR – employees who hold a university degree, a degree by a technical college (Fachhochschule) or who have graduated from a higher vocational school (Höhere Fachschule).

of small enterprises may be viewed as a proxy for the entrepreneurial climate of the region, reflecting the start up activity of previous periods. In contrast, the share of large firms is expected to be negatively related to regional start-ups since a dominance of large plants might create entry barriers (see Armington and Acs 2002, Lee et al. 2004).

Diversity measures

Finally, four different types of diversity are considered in the regression analysis:

Sectoral diversity is probably the most common concept. The indicator of sectoral diversity (*DIV_S*) is based on employment shares of 28 industries.[6]

Technology and entrepreneurship are characterised by an interdependent relationship. On the one hand, technological change is supposed to influence firm birth, on the other hand new firms are considered to be important for the development and diffusion of new products (Verheul et al. 2002). Technology enters the empirical analysis via two ways. Firstly, we aim at investigating determinants of start-ups at different technological levels. Secondly, the significance of technological diversity with respect to new firm foundation is analysed. We approximate technological diversity (*DIV_T*) by analysing the patent portfolio of each region (patent shares of 31 technological areas).

The relevance of cultural diversity might on the one hand rest on more diverse consumer demand that is likely to raise start-ups due to creating market niches and on the other hand on a greater diversity of (complementary) skills that opens up new opportunities for business start-ups. Our indicator measures cultural diversity (*DIV_C*) of the labor force based on regional employment data differentiated by 213 nationalities.

The indicator of occupational diversity (*DIV_O*) is based on regional employment data differentiated by occupation. The applied classification involves a differentiation into 83 occupations. While cultural diversity reflects ethnic and cultural differences between people, occupational diversity is meant to reflect the variety of professional backgrounds of a region's population.

We measure diversity based on two indicators. The first diversity measure is calculated as 1 minus the Herfindahl index of concentration across groups:

$$H_{it} = 1 - \sum_{k=1}^{K} s_{ikt}^{2} \qquad (4)$$

[6] For the corresponding classification of industries see Appendix.

where s_{ikt} is the share of a specific group *k* in the overall sum of the corresponding variable. For instance, in case of cultural diversity s_{ikt} is the share of employees with nationality *k* among all employees of region *i* in year *t*.

The second measure of diversity is the so-called Krugman index. This indicator measures the distribution of a variable relative to a specific reference distribution. Frequently, the average distribution, i.e. for instance the sectoral structure of the country, is used as reference. The Krugman index is calculated as the sum of absolute differences between region-specific shares and shares of the reference distribution $s_{\bullet kt}$:

$$(5) \qquad K_{it} = \sum_{k=1}^{K} \left| s_{ikt} - s_{\bullet kt} \right|$$

Thus, overall we calculate 8 diversity indicators – for each kind of diversity (technological, sectoral, cultural and occupational) a Krugman index and a measure based on the Herfindahl index. In the basic model presented here we restrict ourselves to the 4 Herfindahl measures.

As can be seen from table A2 in the appendix, the correlation between most explanatory variables is relatively low, such that multicollinearity issues should not cause major problems in the regressions.

4. Empirical analysis

4.1 The basic model

The results of the basic model as reflected in equation (1) are given in table 2. As can be seen from table 2, two of the most important determinants of regional start-up intensity (for all kinds of start- ups) appear to be the share of R&D employees (variable RD) and the firm size structure of the region (variables SE and LE). A high percentage of small enterprises (less than 20 employees) in a region is conducive to start-up activity, whereas a high percentage of large enterprises (more than 500 employees) has a dampening effect on start-ups, at least in the high tech sector. A high percentage of small enterprises may be a proxy for the entrepreneurial climate of the region, reflecting the start up activity of previous periods.[7]

R&D employees dispose of a very specific human capital such that people belonging to this group have a particularly high propensity to found new enterprises. Human capital in general, reflected by HQ, the share of highly qualified employees, has a positive impact on start-up activities in all kinds of technology-oriented start ups (which may be due to the fact that technology-oriented start ups are particularly knowledge-intensive) but no significant impact on total start-ups which are mainly not technology oriented. Thus, the results for the two knowledge variables correspond with the theoretical expectations outlined in section 2.

The regions disposable income (INC) has in most cases a positive impact on start-up activity although this variable is rarely significant. The impact is significant at the 10% level only for total start-ups. This might indicate that local demand matters only for new firms that are not technology oriented, whereas high tech firms operate at the European or global level. In contrast, the weighted disposable income of neighbouring regions (W_INC) and it's growth rate (W-INC_GR) have a clearly negative and highly significant impact on start-ups. A possible explanation is that potential founders might prefer to start-up their business in neighbouring regions (move away from the region under consideration) if neighbouring regions offer a large market, as measured by purchasing power, or a particularly dynamic economy.

The unemployment rate UR has a clearly negative impact on regional start-up activities which may be explained by the fact that more prosperous regions offer better conditions for start-ups than problem regions.

[7] We have no direct measure for the average firm age in a region, but there is, of course, a positive relation between firm age and firm size, such that a high percentage of small firms also points to a high percentage of young firms.

Since the share of large firms and the industry density only exert a significant effect in the model for high tech start-ups, they are not considered for the explanation of other firm birth rates. Regarding high tech start-ups the density variable points to the importance of some kind of agglomeration economies linked to the location of industry. In contrast, a firm size structure characterized by a high share of large firms tends to dampen the foundation of high tech firms. This might be due to entry barriers caused by large firms, a phenomenon extensively discussed in the entrepreneurship literature.

Table 2: OLS Regression Results – Pooled Model

	Su_all	Su_ht	Su_tos	Su_to
PD	4.73** (2.54)	0.04** (2.11)	0.19 (1.20)	0.17 (0.67)
INC	1.61* (1.72)	-0.01 (1.20)	0.06 (0.61)	0.13 (0.79)
W_INC	-2.44** (2.11)	-0.02* (1.90)	-0.25** (2.15)	-0.48** (2.46)
INC_GR	39.9 (0.53)	-0.47 (0.47)	2.93 (0.50)	-1.41 (0.15)
W_INC_GR	-263.7** (2.54)	-0.86 (0.66)	-16.7*** (2.67)	-29.8*** (2.91)
UR	-48.0** (2.15)	-0.17 (0.56)	-7.88*** (4.83)	-13.9*** (5.16)
FP	-42.2 (0.73)	-1.02 (1.22)	-5.04 (1.11)	-6.62 (0.82)
HQ	22.7 (0.53)	0.78* (1.90)	12.7*** (3.33)	21.7*** (3.54)
RD	194.2*** (2.64)	1.71** (1.97)	25.0*** (3.49)	37.6*** (3.28)
SE	72.0*** (3.69)	-0.11 (0.29)	6.35*** (3.31)	10.3*** (3.32)
LE	-	-0.59** (2.49)	-	-
ID	-	0.84** (2.17)	-	-
DIV_T	40.5 (1.64)	0.10 (0.43)	1.09 (0.47)	1.49 (0.39)
DIV_C	-17.5 (0.58)	0.80* (1.86)	4.06* (1.87)	8.92** (2.31)
DIV_S	-68.7 (0.85)	-1.16 (0.98)	-5.74 (1.00)	-10.7 (1.04)
DIV_O	-261.7** (2.01)	-0.50 (0.28)	-23.9** (2.11)	-30.6* (1.71)
	R^2=0.28 $R^2_{adj.}$=0.22 F[15,170]= 5.95	R^2=0.47 $R^2_{adj.}$=0.42 F[15,168]=6.24	R^2= =0.67 $R^2_{adj.}$= 0.64 F[15,170]= 21.4	R^2= 0.70 $R^2_{adj.}$=0.67 F[17,168]=25.5

Notes: *** significant at the 0.01 level, ** significant at the 0.05 level, * significant at the 0.10 level. Regions marked by upward bias start-up rates are excluded. All models also include time fixed effects. t-ratios in parentheses are heteroscedasticity consistent.

Impact of diversity

We find no clear-cut pattern in the sense that more diversity is always conducive to start-up activities. There is no significant effect of technological and sectoral diversity on regional firm birth rates. By contrast, occupational diversity (DIV_O) has a significantly negative impact on start-up activities, except for high tech start-ups, whereas cultural diversity

(DIV_C) has a significantly positive impact on technology oriented start-ups in general (SU_TO), technology-oriented services (SU_TOS) and high tech start-ups (SU_HT). So, while there is no important role for the diversity of the regions, heterogeneity of the people seems to matter for entrepreneurial activity. However, the regression results concerning these diversity measures are not always in line with theoretical implications.

In any case, the spatial patterns of firm birth rates differ significantly between high tech and low tech. This implies that it makes sense to estimate separate models for the different groups of start-ups, according to their technology orientation (knowledge intensity). The is reflected by the estimates summarized in Table 2 since there are pronounced differences between the considered start-up rates with respect to the factors that determine the regional disparities.

4.2 Robustness checks: estimation of panel data models

Robustness checks of the pooled regression results include the estimation of panel data models with fixed effects and random effects, i.e. we control for unobserved time-invariant explanatory variables (see equation (2) in chapter 2). In Table 3 we only report estimates for the random effects specification because Hausman tests tend to indicate selection of the random effects model. The results of the Breusch-Pagan test (BP) suggest that there are significant region specific effects.

Comparison of the OLS estimates of the pooled model and the results of the random effects specification shows that, neglecting high tech start-ups, most findings are robust. We detect only a few changes of signs and significance for the dependent variables Su_all, Su_tos and Su_to. These changes mainly refer to the income variables and the diversity measures included in the models. The impact of spatially lagged purchasing power is fairly robust for the sum of technology oriented start-ups and technology oriented services, whereas for total start-ups the significant impact disappears by considering region specific effects. Pooled and GLS estimates suggest some kind of competition effect, i.e. regions with a high disposable income might attract entrepreneurs from neighbouring areas, thus reducing the firm birth rate there. However, there is no longer evidence for a significant influence of the income of the home region in the random effects model. With respect to the growth measures the main findings are largely unchanged, but significance levels are clearly reduced in the GLS model. The prominent role of R&D employment and the firm size structure is underscored by the panel estimates, and the impact of unemployment is even reinforced by taking unobserved time-invariant variables into account. Finally, evidence on the importance of the different forms of diversity for entrepreneurial activity is rather weak in the random effects model. The coefficients of technological and sectoral diversity are not significantly different from zero and there is some indication for dampening effects of occupational diversity.

Table 3: GLS Regression Results – Random Effects Model

	Su_all	Su_ht	Su_tos	Su_to
PD	4.60* (1.95)	0.03 (1.27)	0.24 (1.22)	0.24 (0.72)
INC	1.45 (1.31)	-0.01 (1.07)	0.03 (0.36)	0.09 (0.56)
W_INC	-1.65 (1.18)	-0.02 (1.50)	-0.26** (2.26)	-0.52*** (2.67)
INC_GR	45.0 (0.68)	0.02 (0.03)	7.95 (1.38)	9.51 (0.99)
W_INC_GR	-60.4 (0.68)	-0.28 (0.28)	-14.1* (1.82)	-24.0* (1.86)
UR	-51.5* (1.88)	0.15 (0.47)	-6.70*** (2.92)	-10.1*** (2.62)
FP	-49.8 (0.74)	-0.52 (0.75)	-4.34 (0.77)	-3.59 (0.38)
HQ	36.1 (0.76)	0.70 (1.40)	10.9*** (2.76)	16.8** (2.52)
RD	166.9* (1.78)	1.49 (1.50)	26.6*** (3.40)	42.1*** (3.20)
SE	57.4** (2.22)	0.18 (0.45)	6.19*** (2.88)	10.8*** (2.98)
LE	-	-0.44* (1.95)	-	-
ID	-	0.97** (2.28)	-	-
DIV_T	50.7 (1.40)	0.02 (0.06)	1.72 (0.57)	1.67 (0.33)
DIV_C	-23.5 (0.68)	0.63* (1.67)	3.64 (1.26)	7.69 (1.59)
DIV_S	-87.5 (0.90)	0.28 (0.28)	1.83 (0.22)	10.4 (0.75)
DIV_O	-246.5 (1.46)	-1.60 (0.78)	-29.0** (2.07)	-46.7** (1.98)
BP	32.3***	22.4***	23.2***	20.6***
Hausman test	-[a)]	7.66	12.7	23.1*
	R^2= 0.26 Wald chi2 [15]=39.1	R^2= 0.46 Wald chi2 [17]=175	R^2= 0.66 Wald chi2 [15]=220	R^2= 0.69 Wald chi2 [15]=237

Notes: *** significant at the 0.01 level, ** significant at the 0.05 level, * significant at the 0.10 level. Regions marked by upward bias start-up rates are excluded. All models also include time fixed effects.
a) Model fails to meet asymptotic assumptions of Hausman test.

More pronounced changes emerge with respect to high tech start-ups. The majority of significant effects detected in the pooled model vanish if we allow for region specific effects. There are only three influential factors left in the random effects model. A firm size structure characterized by a relatively high share of large firms tends to dampen entrepreneurial activity, whereas high industry density and cultural diversity of the work force seem to be important components of an environment that is conducive to high tech start-ups. In contrast, the positive effects of R&D activity and human capital identified in the pooled model do not differ from zero at conventional levels of significance if we apply the panel approach.

4.3 Systematic differences between East and West?

Finally, we investigate whether there are still systematic differences in entrepreneurial activity between East and West German regions. The weak performance of some random effects models, especially regarding high tech start-ups, might at least partly be caused by the assumed homogeneity of the slope parameters. Spatial regimes models are applied to deal with this issue, i.e. we allow for different intercepts and slopes for the two subsets of the data (Anselin 1988). We analyse the stability of the coefficients over the two regimes (East, West) via Wald tests on the equality of slopes estimated for the subsets. The results of the Wald tests are summarized in Table 4. The corresponding coefficient estimates are given in Table 5. All results refer to random effects specifications.

Table 4: Wald Test on Equality of Slope Estimates

	Su_all	Su_ht	Su_tos	Su_to
PD	0.17	0.58	0.14	1.66
INC	0.58	0.06	0.01	0.38
W_INC	1.43	4.79**	1.39	3.14*
INC_GR	5.46**	0.35	4.83**	3.24*
W_INC_GR	3.20*	0.29	8.47***	8.75***
UR	1.14	2.54	0.11	1.07
FP	0.56	1.74	0.27	0.98
HQ	1.24	8.18***	0.63	1.68
RD	0.64	6.55**	0.06	0.02
SE	2.60	0.92	1.51	2.52
LE	0.65	1.77	0.29	0.19
ID	6.31**	6.86***	4.35**	5.74**
DIV_T	0.66	0.00	1.37	0.78
DIV_C	3.98**	0.00	0.04	0.15
DIV_S	1.52	1.29	0.00	0.20
DIV_O	7.83***	0.08	2.94*	3.27*

Notes: The Wald statistic given in the table has an asymptotic chi-square distribution with one degree of freedom.
*** significant at the 0.01 level, ** significant at the 0.05 level, * significant at the 0.10 level.
Regions marked by upward bias start-up rates are excluded.

The coefficient estimates indicate that both parts of the country are still characterized by significant differences regarding the factors that influence the regional variation of start-ups.

However, disparities do not pertain to all variables. Especially the impact of income growth, growth in neighbouring regions, industry density and occupational diversity differs between the former communist part and West Germany. Whereas the high income growth has an adverse effect on entrepreneurial activity in West German regions, start-ups tend to benefit from growth in the Eastern part of the country. The negative coefficients for the Western subset might point to high opportunity costs of starting a business in regions marked by a dynamic development because of attractive job offers. Since unemployment is generally much higher in East Germany this mechanism might not work. In spite of high growth the labour market situation continues to be tense.

Pronounced differences also arise for industry density. The coefficient of the variable is positive, though not significant in most models, for the West German cross section. In contrast, a relatively high industry density seems dampen firm foundation in East Germany. This finding might be caused by the structural disparities between both parts of the country. Average industry density in West Germany almost doubles the level in the Eastern part. Locations in East Germany marked by a relatively high industry density are perhaps still dominated by some successor of the large production cooperatives of the former communist economy. Thus the negative effect might indicate entry barriers caused by the dominance of large plants that are not relevant in West German regions.

Finally, occupational diversity seems to matter only in West German regions. Contrary to theoretical expectations the impact is negative. Significant differences regarding the impact of cultural diversity only emerge for the overall firm birth rate with a positive effect in the East whereas in West Germany the variable exerts no prominent influence. At first sight, the coefficient estimates also suggest that there might be systematic differences with respect to unemployment because significant negative effects are restricted to the East German subset. Moreover, we might expect corresponding evidence owing to the pronounced labour market disparities. However, as indicated by the Wald tests, the variations regarding this explanatory variable are not statistically significant.

The evidence described so far mainly refers to total start-ups, technology oriented firms and technology oriented services. For high tech start-ups we arrive at a specific pattern. Income growth is no source of differentiation with respect to new high tech firms because it does not matter in both parts of the country. However, for industry density the familiar result arises. Besides, differences primarily apply to factors that are supposed to play a prominent role in promoting high tech start-ups, namely human capital and R&D activity. The effect of the share of highly skilled workers on entrepreneurial activity is not significantly different from zero for the Western part of the country, whereas the impact is positive in East German regions. In contrast, in West Germany it is rather R&D activity than the availability of human capital that exerts a positive effect on the foundation of high tech businesses, while the impact of R&D employment in East Germany is negative. This finding might indicate some kind of threshold effect being at work. In most East German regions the overall level of R&D might be too low for significant spillover effects that could promote start-ups of innovative firms. In fact, the average share of R&D employment in East Germany was 1.4 percentage points below the West German level in 2001.

Altogether, the evidence indicates that significant differences between East and West German regions regarding the determinants of regional firm birth rates characterize all levels of technology. However, disparities among East and West differ between new high tech firms and other start-ups. The variations among both parts of the country seem to be driven by the pronounced structural differences and perhaps some kind of threshold effects. Moreover, some regression results seem to be severely biased by not taking account of the structural disparities within the country. This refers for instance to the negative impact of the share of large firms that emerges under the assumption of homogenous slope parameters. There is no evidence for a negative effect once we allow the coefficient to differ between the subsets of regions. In fact, the findings of the spatial regimes model rather indicate that large firms might exert a positive influence on the regional firm birth rate. Moreover, we detect significant positive effects of sectoral diversity applying the spatial regimes models for technology oriented services and the sum of technology oriented start-ups that did not show up in the other regression models.

Table 5: Regression Results – Spatial Regimes Model

	Su_all		Su_ht		Su_tos		Su_to	
	East	West	East	West	East	West	East	West
PD	3.10	7.55*	0.21	0.09	0.32	-0.05	1.87	-0.11
INC	-4.69	1.25	-0.03	-0.01	-0.04	0.02	-0.48	0.07
W_INC	0.93	-3.39**	0.08	-0.04**	0.04	-0.34*	0.34	-0.63*
INC_GR	127.5	-192.2**	1.90	0.69	17.0**	-5.12	29.6**	-3.30
W_INC_GR	1.15	-312.9**	1.11	-0.18	-0.26	-39.4***	8.36	-61.2***
UR	-106.1**	-5.70	-0.96**	0.39	-8.64***	-6.28	-17.4***	-4.93
FP	-276.7	-1.53	-7.87	-0.47	-6.62	7.54	-29.4	13.7
HQ	71.9	-87.9	4.25***	-0.38	6.28	-3.72	16.5	-9.61
RD	-29.4	179.6*	-7.28**	2.31*	30.0	24.6**	37.0	42.2***
SE	17.4	164.0***	-0.51	0.53	8.03*	15.6***	12.5*	29.0***
LE	77.7*	37.0*	0.82	-0.41	3.24	5.71***	6.47	9.86***
ID	-261.0**	101.5	-2.54**	1.17**	-12.3*	6.82	-23.7*	16.6
DIV_T	98.1	30.6	-0.08	-0.05	8.28	0.66	9.62	0.10
DIV_C	361.5*	-16.2	0.70	0.70	4.80	1.74	-5.04	4.54
DIV_S	-91.3	193.0	4.93	-0.72	19.8	20.0***	49.7**	35.8**
DIV_O	567.3	-1012.8	-0.61	-2.74	-17.9	-81.6***	-34.7	-155.9***
	R^2= 0,43 Wald chi2 [35]= 92.0		R^2= 0,55 Wald chi2 [35]= 206.7		R^2= 0,68 Wald chi2 [35]= 234.3		R^2= 0,70 Wald chi2 [35]= 234.8	

Notes: The Wald statistic given in the table has an asymptotic chi-square distribution with one degree of freedom.
*** significant at the 0.01 level, ** significant at the 0.05 level, * significant at the 0.10 level.
Regions marked by upward bias start-up rates are excluded.

5. Conclusions and Outlook

In this paper we have analysed the regional determinants of entrepreneurship in Germany, giving special emphasis to regional diversity in its various dimensions.

We find that the spatial patterns of firm birth rates differ significantly with respect to the technology (or knowledge) level necessary to start up a certain business and that there are still systematic differences between the determinants of firm birth between the (former communist) eastern and western part of Germany. The most important determinants of regional start-up intensity appear to be knowledge capital (as measured in terms of R&D employment), firm size structure, weighted disposable income and the regional unemployment rate which has a clearly negative impact on firm start-ups. As concerns the impact of diversity we find no clear-cut pattern in the sense that more diversity is always (and in all its dimensions) conducive to start-up activities. Occupational diversity tends to have a negative impact on start-up activity whereas cultural diversity appears to be conducive to firm births, in particular in the more technology- (knowledge-) intensive sectors of the economy. Our results are, however, still preliminary, in particular with respect to the impact of diversity on start-ups.

We intend to extend the regression model by allowing for interaction effects, i.e. combining the heterogeneity of the people (measured by cultural and occupational diversity) with the regional stock of knowledge, R&D employment, because modelling the impact of diversity in this way might be more appropriate than analysing the effects of knowledge and diversity separately. Moreover, it is planned to investigate in more detail the mechanisms behind the East-West differences.

An important issue for future research might be to check systematically for neighbourhood effects via considering spatial lags not only for the income variables but also for the knowledge related variables since we have virtually no information on the spatial range of knowledge spillovers.

Appendix

Table A1 Industry classification

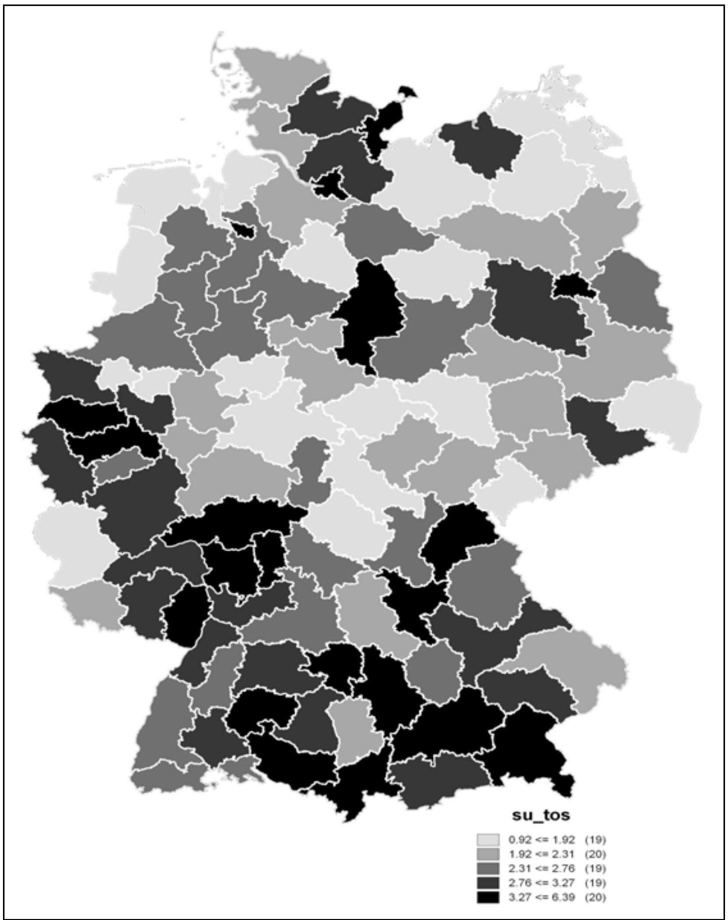
The indicator for sectoral diversity is based on employment data by region and industry. The following classification is applied:

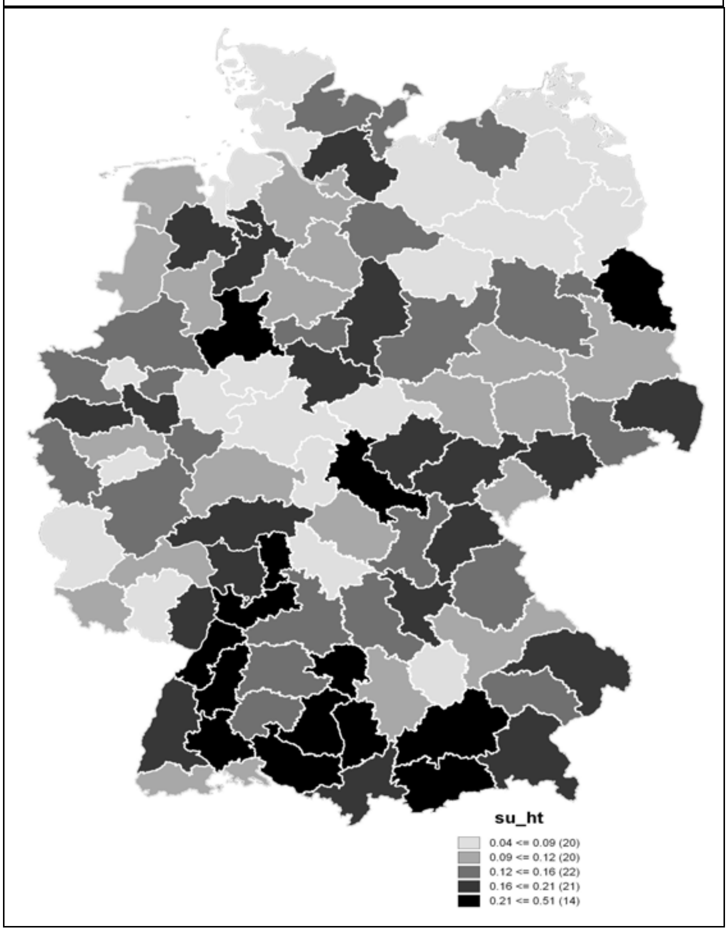
1. Agriculture, hunting and forestry
2. Energy
3. Mining
4. Chemical industry
5. Rubber and plastic products
6. Non-metallic mineral mining
7. Glass and ceramics
8. Basic metals and fabricated metal products
9. Machinery
10. Transport equipment
11. Electrical and optical equipment
12. Manufacturing n.e.c.
13. Wood and wood products
14. Pulp, paper and paper products, publishing and printing
15. Leather and textiles
16. Food, beverages and tobacco
17. Construction
18. Wholesale and retail trade
19. Transport and communication
20. Financial intermediation
21. Hotels and restaurants
22. Health and social work
23. Business services
24. Education
25. Leisure-related services
26. Household-related services
27. Social services
28. Public sector

Table A2: Correlation analysis – explanatory variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	PD	1.00																			
2	ID	-0.06	1.00																		
3	INC	0.63	0.10	1.00																	
4	W_INC	0.10	0.22	0.24	1.00																
5	INC_G	-0.27	0.11	-0.08	0.15	1.00															
6	W_INC_G	0.13	0.31	0.15	0.01	0.18	1.00														
7	UR	0.06	-0.68	-0.17	-0.36	-0.26	-0.50	1.00													
8	FP	0.52	0.48	0.66	0.39	-0.01	0.28	-0.57	1.00												
9	HQ	0.35	-0.28	0.47	-0.09	0.10	-0.08	0.38	0.15	1.00											
10	RD	0.33	0.45	0.56	0.28	-0.00	0.17	-0.37	0.74	0.43	1.00										
11	SE	-0.49	-0.40	-0.57	-0.18	0.07	-0.08	0.13	-0.63	-0.42	-0.71	1.00									
12	LE	0.54	0.34	0.61	0.20	-0.12	0.15	-0.26	0.71	0.33	0.77	-0.85	1.00								
13	DIV_T_H	0.13	-0.10	0.20	0.15	0.05	0.03	-0.03	0.19	0.16	0.02	0.03	-0.07	1.00							
14	DIV_T_K	-0.15	-0.25	-0.29	-0.23	-0.08	-0.20	0.38	-0.43	-0.13	-0.26	0.18	-0.22	-0.65	1.00						
15	DIV_C_H	0.30	0.60	0.56	0.39	0.06	0.34	-0.69	0.93	0.02	0.70	-0.46	0.58	0.21	-0.48	1.00					
16	DIV_C_K	-0.33	-0.53	-0.57	-0.43	-0.07	-0.34	0.77	-0.92	0.01	-0.66	0.44	-0.59	-0.19	0.46	-0.94	1.00				
17	DIV_S_H	-0.21	0.34	-0.14	0.16	-0.10	0.07	-0.30	0.03	-0.22	0.00	0.11	-0.17	0.17	-0.33	0.17	-0.11	1.00			
18	DIV_S_K	-0.05	-0.04	-0.21	-0.23	0.06	-0.10	0.36	-0.27	-0.00	-0.23	0.12	-0.15	-0.20	0.39	-0.27	0.35	-0.48	1.00		
19	DIV_O_H	-0.47	0.19	-0.58	-0.16	0.03	-0.11	0.19	-0.55	-0.32	-0.36	0.38	-0.46	-0.24	0.23	-0.40	0.53	0.32	0.18	1.00	
20	DIV_O_K	-0.11	-0.23	-0.25	-0.29	0.11	-0.12	0.41	-0.40	0.03	-0.42	0.33	-0.35	-0.18	0.38	-0.41	0.43	-0.35	0.80	0.25	1.00

Map A1: Regional distribution of technology-oriented services and high-tech start-ups





References

Acs, Z., Audretsch, D., Braunerhjelm, P. and Carlsson, B. (2004). *The Missing Link: The Knowledge Filter and Endogenous Growth* (Discussion paper). Stockholm: Center for Business and Policy Studies.

Anselin, L. (1988): Spatial Econometrics, Methods and Models. Dordrecht, Kluwer Academic.

Armington, C. and Acs, Z.J. (2002): The Determinants of Regional Variation in New Firm Formation, *Regional Studies* 36 (1): 33-45.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity*. Princeton: Princeton University Press, 609-626.

Audretsch, David B. and Dirk Dohse (2007), Location: A Neglected Determinant of Firm Growth. *Review of World Economics* 143 (1): 79-107.

Audretsch, David B. and Max Keilbach, (forthcoming) "The Theory of Knowledge Spillover Entrepreneurship," *Journal of Management Studies*.

Audretsch, David B., Max Keilbach and Erik Lehmann (2006). *Entrepreneurship and Economic Growth*, New York: Oxford University Press.

BBR (Bundesamt für Bauwesen und Raumordnung) (2001). Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden. Berichte. Band 8: Bonn.

Carlton, D. (1983). The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. *Review of Economics and Statistics, 54*, 440-449.

Feldman, M. and Audretsch, D. (1999). Innovations in Cities: Science-Based Diversity, Specialization and Localized Monopoly. *European Economic Review, 43*, 409-429.

Glaeser, E., Kallal H., Scheinkman J. and Shleifer A. (1992). Growth in Cities. *Journal of Political Economy*, *100*, 1126-1152.

Jacobs, J. (1969). *The Economy of Cities*. New York: Vintage Books.

Jaffe, A. (1989). The Real Effects of Academic Research. *American Economic Review*, *79,* 957-970.

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics, 63*, 577-598.

Lee, S.Y, R. Florida and Z. Acs (2004): Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation, *Regional Studies* 38 (8): 879-891.

Metzger, G.; Heger, D. (2005): Die Bereitstellung von Standardauswertungen zum Gründungsgeschehen in Deutschland und Österreich für externe Datennutzer, Version 2005-02, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim.

Porter, M. (1990). *The Comparative Advantage of Nations*. New York: Free Press.

Reynolds, P., Storey, D. and Westhead, P. (1994). "Regional Variations in New Firm Formation – Special Issue,". *Regional Studies, 28*, 343-456.

Rosenthal, S. and W. Strange (2003). Geography, Industrial Organizations and Agglomeration. *Review of Economics and Statistics* 85 (2): 377-393.

Stuart, T. and O. Sorenson (2003). The Geography of Opportunity: Spatial Heterogeneity in Founding Rates and the Performance of Biotechnology Firms. *Research Policy*, 32, 229-253.

Storey, D. (1991). The Birth of New Firms – Does Unemployment Matter? A Review of the Evidence. *Small Business Economics, 3,* 167-178.

Storey, D.J. (1994): Understanding the Small Business Sector, Routledge, London/New York.

Verheul, I.; Wennekers, S.; Audretsch, D. and R. Thurik (2002): An eclectic theory of entrepreneurship: policies, institutions and cultures, in: Audretsch, D.; Thurik, R.; Verheul, I. and S. Wennekers (eds.), Entrepreneurship: determinants and policy in a European-US comparison, Kluwer, Boston Mass., pp. 11-81