Organizational Research Methods
Volume 13 Number 1
January 2010 176-191
© 2010 SAGE Publications
10.1177/1094428109338871
http://orm.sagepub.com
hosted at
http://online.sagepub.com

Dealing With Spatial Heterogeneity in Entrepreneurship Research

Robert J. Breitenecker
Klagenfurt University
Rainer Harms
University of Twente

In quantitative research, analyses are generally made using a geographically defined population as the study area. In this context, the relationships between predictor and response variables can differ within the study area, a feature that is known as spatial heterogeneity. Without analyzing spatial heterogeneity, a global model may not be correct, and there may be unclear spatial boundaries in the generalizability of the findings. The authors discuss how the method of geographically weighted regression (GWR) can be used to identify the study area, and illustrate the utility of GWR for empirical analyses in entrepreneurship research. Future entrepreneurship research can benefit from analyzing whether conflicting evidence may be due to spatial heterogeneity and from applying GWR in an exploratory way.

Keywords: *spatial heterogeneity; entrepreneurship; start up rate; geographically weighted regression*

In entrepreneurship research, spatial aspects can play an important role in many analyses, such as the diffusion of innovations (Rogers, 1983), market entry of firms (Debarsy & Dejardin, 2008), and the role of regional characteristics in shaping entrepreneurial activity (Karlsson & Dahlberg, 2003). However, even though georeferenced data and suitable software packages are becoming increasingly available (Anselin & Florax, 1995), the field has not yet embraced methods for spatial analysis. For example, researchers that analyze an entrepreneurial phenomenon in a geographically defined population in a specific country may implicitly assume that the results are the same in all parts of that study area. However, many researchers do not currently use statistical tools to support this homogeneity assumption.

If this homogeneity assumption is violated, the relationships between predictor variables and a response variable vary between different local study areas, which create spatial heterogeneity (Anselin, 1988). When spatial heterogeneity exists, analyses can be influenced by counterbalancing (e.g., effects that cancel each other out) or dilution (e.g., averaging different size effects). Therefore, a model that assumes constant parameters across the study area would not be valid. By using local spatial analysis as a "spatial microscope"

Authors' Note: Thanks to special issue editors Jim Combs, Duane Ireland, Dave Ketchen, Jeremy Short, and three anonymous reviewers. Thanks to Lori Peterson for her assistance in improving this manuscript. Please address correspondence to Robert J. Breitenecker, Department of Innovation Management and Entrepreneurship, Klagenfurt University, Universitätsstr. 65-67, A-9020 Klagenfurt, Austria; e-mail: robert.breitenecker@uni-klu.ac.at.

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

Table 1
Occurrence and Analysis of Spatial Heterogeneity

	Spatial Heterogeneity		
	Not Analyzed	Analyzed and Not Discovered	Analyzed and Discovered
Spatial heterogeneity exists	Leads to erroneous results	Type II error	Methods from spatial econometrics can be applied
Spatial heterogeneity does not exist	Regional boundaries of analysis unclear	Make a stronger claim for generalization	Type I error

(Fotheringham, Brunsdon, & Charlton, 2002, p. 252), researchers could investigate the homogeneity respectively the heterogeneity of their study area.

In this article, we demonstrate how entrepreneurship research can benefit from recognizing spatial heterogeneity. First, we discuss the importance of including spatial heterogeneity in entrepreneurship research through an analysis of the literature. Then, we introduce the geographically weighted regression (GWR) as a tool to detect spatial heterogeneity. An illustrative analysis demonstrates how by investigating spatial heterogeneity, counterbalancing and dilution effects can be avoided. Researchers who use this method may help to reconcile conflicting evidence in entrepreneurship research.

Addressing Spatial Heterogeneity in Entrepreneurship Research

Consequences of Inadequately Dealing With Spatially Heterogeneous Study Areas

Entrepreneurship research, as a discipline, analyzes how the process of venture creation is shaped by characteristics of the person, of the venture, and of the environment (Davidsson, 2005). A key element of the environment is the geographical location, because location influences the process and the results of venture creation. For example, studies on cluster dynamics and systems of innovation (Carayannis, Assimakopoulos, & Kodo, 2008) and the local dynamics of entrepreneurship (Julien, 2007) highlight the location-specific nature of entrepreneurship.

This location-specific nature of entrepreneurship can be understood as a "*spatial conditionality*," which describes how features of the environment can influence the strength and/or the direction of the relationship between a predictor variable and a response variable. Although contingent relationships in entrepreneurship research at the industry level have frequently been found (Lumpkin & Dess, 1996), it is likely that heterogeneity also exists at a regional level, because industry contingency factors may be unequally distributed in space.

Table 1 illustrates the consequences of spatial heterogeneity in quantitative research. The table indicates whether spatial heterogeneity does or does not exist in the observable environment and if it has been analyzed.

If spatial heterogeneity exists, but it is not analyzed, then the implicit assumption of the researcher is that spatial homogeneity exists, an assumption that can lead to erroneous results. For instance, positive and negative relationships between predictor and response

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

variables in local study areas may cancel each other out, creating a *counterbalancing effect*. A weaker form of counterbalancing effects arises when variations in the strengths of relationships are averaged over a study area, without a change in direction, producing a *dilution effect*. Counterbalancing and dilution effects may be strongest when no spatial differentiation of a heterogeneous study area is undertaken. However, these effects may also emerge in finer-grained analyses when the study areas are not specifically based on spatial heterogeneity.

If spatial heterogeneity does not exist and is not part of the selection criteria either then the *spatial boundaries* of the analysis *remain unclear*. In these circumstances, the researcher can only imply that the findings hold true for the whole study area but cannot demonstrate it. When spatial heterogeneity exists but is not detected by an adequate analysis, a type II error is made (rejecting the true null hypothesis of no spatial variation). When a dedicated analysis of spatial heterogeneity leads to no results, a stronger claim for the generalizability of the findings over the study area can be made. If spatial heterogeneity exists and is discovered then dedicated techniques of spatial econometrics are appropriate. Techniques such as spatially switching regression models (Anselin, 1988) can be used to analyze the relationship of the predictor variables with the response variable in a spatial context. If spatial heterogeneity does not exist but the spatial heterogeneity analysis leads to an erroneous result of spatial varying local associations between predictor and response variable, a type I error occurs (retaining the false null hypothesis of no spatial variation).

Potential Approaches to Deal With Spatial Heterogeneity

By delineating spatially homogenous study areas, researchers can avoid counterbalancing and dilution effects and can provide information on the geographic scope of the generalizability of their findings. Now, the question arises as how researchers can actually identify those homogenous study areas. As a starting point for spatial analyses, researchers are likely to use official statistics describing administrative statistical units such as Nomenclature of Territorial Units for Statistics (NUTS, Eurostat, 2009) and select local study areas within the broader study area *ex ante*. In principle, the inclusion of interaction effects with dummy variables that represent those regions could solve the issue of spatial heterogeneity. If the interaction term with the regional dummy variable were significant, this could be interpreted as a change in the slope of the relationship between predictor and response variables, thereby modeling structurally different relationships. Another strategy could be to run several regression analyses, each of which is based on data from a specific local study area, and compare the results. Differences between the regressions in terms of signs and values of the regression coefficients would indicate structurally heterogeneous relationships.

However, the *ex ante* selection of the local study areas is exacerbated by: (a) the large potential number of local study areas, (b) limited theoretical knowledge, and (c) potential interdependencies between study areas. These issues, as discussed in the following paragraphs, may persuade researchers of the need to apply methods specifically suited to identify spatial heterogeneity.

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

First, the number of regional units per country can be quite large. For example, there are 93 study areas at the local area unit (LAU) level in our analysis in Austria. For researchers who wish to assess spatial heterogeneity in a national context via the inclusion of regional dummy variables in moderated regression, this number of units would result in a model with a high degree of complexity and a low degree of freedom, which would be difficult to interpret.

Second, as an alternative to resorting to regional dummy variables, researchers could include regional-specific variables as moderator variables. For example, researchers could use the unemployment rate or regional income directly as moderator variables. However, without first testing whether spatial heterogeneity is present, the results would be unclear. A concern to researchers is the comprehensiveness of the incorporated moderators and the degree to which spatial variability in total can be explained. Although this approach would be suitable to identify heterogeneous regions, the main deterrent is the need for a theoretical justification in choosing moderator variables. Because the impact of the environment on the entrepreneurial process is complex (Julien, 2007) and not well explored, many regional contingencies have yet to be fully developed from a theoretical perspective.

Third, when data are sampled from a specific local study area, the recognized characteristics of adjacent regions can also influence the relationships between predictor and response variables of that region. This phenomenon may influence inference statistics and cause the significance of the results to be reported incorrectly. The influence of adjacent regions is particularly relevant in entrepreneurship research. Because entrepreneurs are characterized by innovative and proactive behavior (Miller, 1983), their actions may not only be influenced by the resources and opportunities in their ''home regions'' but also by the resources and opportunities of neighboring regions as well. For example, rural entrepreneurs may do business in adjacent urban areas. If the impact of the neighboring urban region on entrepreneurial behavior is stronger than the impact of the rural region's own characteristics, then an *ex ante* characterization as ''rural'' would not match the true behaviour of enterpreneurs in that region.

A Literature Analysis on the Recognition of Spatial Heterogeneity in Entrepreneurship Research

To assess how spatial heterogeneity is recognized in entrepreneurship research, we investigated quantitative studies that analyze determinants of new firm formation. New firm formation is the number or the proportion of new firms that are created in a given area during a given time. We acknowledge that these studies cover only a facet of entrepreneurship research. However, because the response variable can only be operationalized in terms of geographic boundaries and data are often only available on a specific regional level, we believe that these studies are particularly likely to incorporate spatial heterogeneity.

As a starting point for our literature analysis, we used key word searches for studies with ''*new firm formation*,'' ''*new business formation*,'' ''*new venture formation*,'' ''*start-up activity*,'' and ''*firm birth*'' in the titles as well as in the abstracts of five leading entrepreneurship journals (*Entrepreneurship and Regional Development, Entrepreneurship Theory and Practice, Journal of Business Venturing, Journal of Small Business Management, Small Business Economics*) from 1998 to 2008. Additionally, journals from the field of regional science such as ''*Regional Science*'' and ''*Annals of Regional Science*'' were

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

searched and relevant articles were included. We then searched the reference sections of the identified articles for additional studies. The search resulted in 40 articles from 14 journals analyzing the relationship between regional predictor variables and the rate of new firm formation.[1]

The resulting 40 studies exhibited one of three approaches to recognize spatial heterogeneity: (a) an *implicit* approach, (b) a *level* approach, and (c) a *structural* approach. Twenty articles defined the population at a national or regional level. As such, this practice *implied* that the results might be different in areas other than the one studied, but there is no variation within the study area. In more than a third of the studies, researchers recognized that location might influence the *level* of the response variable. For example, the urban incubator hypothesis states that the level of start-up activity may be higher in urban areas than in rural areas (Tödling & Wanzenböck, 2003). When using this hypothesis, researchers may account for this systematic difference by using a spatial dummy variable (Brixy & Grotz, 2007).

In only three studies (Fritsch, 2004; Reynolds, Storey, & Westhead, 1994; van Oort & Atzema, 2004), the researchers recognized that the location affects the *structure* of the relationship between predictor and response variables. These researchers indicated there may be spatial heterogeneity involved in their studies. When considering the results, the three studies showed how the recognition of spatial heterogeneity can lead to more fine grained results.

Recognizing there may be spatial heterogeneity between countries (Reynolds et al., 1994) or between parts of a country (Fritsch, 2004), methodologically, these researchers defined local study areas *ex ante* and calculated separate regressions for local study areas. Both studies showed that the strength of the relationship between predictor variables and new firm formation differs between local study areas, thereby providing evidence for dilution effects.

Van Oort and Atzema (2004) analyzed how the impact of predictor variables on new firm formation differs between local study areas such as macrozoning regimes (e.g., core, intermediate, and peripheral areas of national economic activity) or connected and unconnected regimes (areas that are connected/unconnected by commuter streams). By comparing separate regressions, they showed that the results differed depending on the local study area chosen. In contrast to the other two papers, van Oort and Atzema (2004) introduced a spatial lag model that takes into account spatial dependence in the estimation of the model parameters. They also introduced test statistics, using a spatial Chow-Wald test, to demonstrate that the local study areas identified are indeed structurally different. Overall, however, our review of the literature shows that most studies do not currently deal with spatial heterogeneity.

Using GWR to Detect Spatial Heterogeneity

The recognition of spatial heterogeneity is promising for entrepreneurship research. However, currently most studies do not attempt to identify structurally heterogeneous study areas. In this section, we illustrate GWR, a technique that can be used to detect spatial heterogeneity. We outline the statistical foundations and apply GWR on an example drawn

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

from entrepreneurship research. In the discussion and in the example, we follow a three-step approach: (a) GWR model estimation; (b) calculation of parameter variability; and (c) analysis of relationship stability with a Monte-Carlo test.

Estimating the GWR Model

Prior to GWR analysis, it needs to be assessed whether a key assumption of a global model, that is spatially independent error terms, is violated. If this is the case (Moran *I* statistics, Doh & Hahn, 2008), spatial autocorrelation exists, and that may be due to spatial heterogeneity (Brunsdon, Fotheringham, & Charlton, 1999). Hence, a GWR model must be estimated. As an extension of the classical linear regression model, GWR considers the spatial variation of the relationships between response and predictor variables as well as their spatial dependence. In a GWR model, the parameters are allowed to vary over the entire study area, while being estimated on a local level. Therefore, GWR presents a local model where the parameters are not restricted to one level, as is accepted practice in classical global models. Within the GWR approach, we estimate the association between response and predictor variables for each region of the study area, as such, the GWR can be used as an explorative tool to detect spatial variability (Fotheringham, Brunsdon, & Charlton, 2000). When the variables do not demonstrate spatial variability over the whole study area, a global model would be valid. For other variables, a localized regional structural variable has to be identified. By incorporating interaction effects with these structural variables into the model, the problem of spatial heterogeneity may be solved. The GWR model can be written as

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{K} \beta_k(u_i, v_i)x_{ik} + \epsilon_i, \tag{1}$$

where $(y_i; x_{i1}, \ldots, x_{iK})$ are observations in region i $(i = 1, \ldots, N)$ of the response variable y and predictor variables $x_1, \ldots, x_K,$ and $\beta_k\ (u_i, v_i) = \beta_{ik}$ are the parameters that vary in each region i and k $(k = 0, \ldots, K)$. The regional values are represented by the two dimensional coordinates, such as the longitude and the latitude (u_i, v_i) of its centroids (Fotheringham et al., 2000).

In spatial data, it cannot be assumed that observations are spatially independent. Regions that are closer to each other may have more in common than regions that are more distant; therefore, the observations are spatially dependent (Doh & Hahn, 2008). Tobler (1970, p. 236) has formulated this spatial phenomenon in the first law of geography: “Everything is related to everything else, but near things are more related than distant things.” If spatial dependence is present, it has to be considered in a regression model. GWR considers spatial dependence of regions in the estimation process. It is estimated by weighted least squares and the estimator $\hat{\boldsymbol{\beta}}_i = \left(\hat{\beta}_{i0}, \hat{\beta}_{i1}, \ldots, \hat{\beta}_{iK}\right)^{\mathrm{T}}$ of $\boldsymbol{\beta}_i = (\beta_{i0}, \beta_{i1}, \ldots, \beta_{iK})^{\mathrm{T}}$ for region i is given by

$$\hat{\boldsymbol{\beta}}_i = \left(\mathbf{X}^T\mathbf{W}_i\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}_i y. \tag{2}$$

The spatial weights $w_j^{(i)}$ of the diagonal matrix $\mathbf{W}_i$ are calculated from a spatial kernel function. By estimating the local model of region i, a region j that is closer to region i is given more weight than a region that is farther away. In our illustrative example in the next section, we use a bisquare kernel. This kernel function is written as

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

$$w_j^{(i)} = \begin{cases} (1 - d_{ij}^2/h^2)^2 & \text{if } d_{ij} < h \\ 0 & \text{else} \end{cases}, \tag{3}$$

where d_{ij} is the Euclidian distance between the centroids of region i and region j (Fotheringham et al., 2000). In the local GWR estimation procedure of region i, only regions j inside a predefined radius or bandwidth h ($d_{ij} < h$) are included. Thus, the bandwidth h determines how many neighboring regions should be considered for the local estimate. The bandwidth can be fixed or can vary with location i. When a fixed bandwidth h is used, the number of observations that are considered for the local model varies with each region. An alternative is an adaptive kernel function with varying bandwidth h_i. In this case, the number of considered regions in the local estimates is held constant. The optimal bandwidth h_i can be found by minimization of the Akaike Information Criteria (AIC; Brunsdon, Fotheringham, & Charlton, 1998).

Determining Parameter Variability

Second, once the local GWR parameters have been estimated, summary statistics (minimum, maximum, median) of the local parameter estimates can be used to check whether sign changes of parameter values over the whole study area exist. The variability of the local parameter estimates can be measured by the standard deviations of the local parameter estimates $\hat{\beta}_{ik}$:

$$SD_k = \sqrt{\sum_i (\hat{\beta}_{ik} - \hat{\beta}_k)^2 / N} \tag{4}$$

where $\hat{\beta}_k$ is the mean of the N local parameter estimates of variable k including the intercept ($k = 0, \ldots, K$). The smaller the standard deviation SD_k, the stronger is the evidence of a constant influence of the parameter k over the whole study area. Brunsdon et al. (1998) suggest comparing the standard deviations SD_k with the standard errors of parameter estimates of the global model to determine the degree of spatial variability. If the standard deviations exceed the standard errors, there is evidence of spatial heterogeneity.

Testing Spatial Stability

Third, the significance of spatial variation in the relationships between response and predictor variables must be analyzed. To test for spatial instability of a global model, parametric tests (Brunsdon et al., 1999; Fotheringham et al., 2002) or independent Monte-Carlo tests can be applied.

The Monte-Carlo test begins with the null hypotheses that a global model is valid and that the local parameter estimates $\hat{\beta}_{ik}$ do not vary with the location i for variable k. A set of L GWR estimations (e.g., $L = 999$ permutations) with permuted attribute values over the study area are performed. For each permuted estimation, the standard deviation SD_{kl} ($k = 0, \ldots, K; l = 0, \ldots, L$) is computed (Fotheringham et al., 2002). The set of L standard deviations SD_{kl} for each of the ($K + 1$) local parameter sets generates a sampling distribution for the parameter variability of variable k. If the standard deviation SD_{k0} from the original model (the correctly located case) is an extreme value of this sampling distribution ($p <$

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

.05, one-sided test), then the hypotheses of a constant parameter over the whole study area has to be rejected (Fotheringham et al., 2002).

A *p* value can be calculated under the null hypothesis that the local parameter does not vary over the study area. The *p* value is then the proportion of standard deviations SD_{kl}, which are higher (more extreme) than the standard deviation SD_{k0} of the correctly located case out of all SD_{kl}. Hence, it is given by $1 - R/(L + 1)$, where R is the rank of SD_{k0} out of all SD_{kl} and L the number of permutations.

An Illustration

Motivation and Global Model Estimation

In entrepreneurship research, the analysis of factors that influence the annual number of new start-ups is of great importance for both researchers and policy makers (Audretsch & Fritsch, 1994). By understanding the factors that explain start-up rates, policy makers are able to compare regions and decide on policies to increase start-up activity. Table 2 provides an illustrative overview of the factors that we have used in the analysis that follows.

Table 2 demonstrates how the theoretical discussion and the empirical evidence on almost every predictor variable point in different directions. A possible way to reconcile the conflicting theoretical arguments and empirical evidence would be to identify whether there are regions where one or the other hypothesis holds. This approach will avoid counterbalancing and dilution effects, while at the same time assessing the generalizablity of the findings. First, local study areas that are homogeneous with regard to a relationship under investigation need to be identified and analyzed to discover the underlying causes of potentially diverging relationships.

Table 2
Factors That Affect New Firm Formation: Conflicting Findings

Factor	Explanation	Relationship	Authors
Exit rate	Exiting firms open competitive space for newcomers	+	Anyadike-Danes, Hart, and O'Reilly (2005)
Unemployment	Self-employment as only alternative to generate income	+	Ritsilä and Tervo (2002)
	Self-employment as chance to profit from opportunities	–	
Average net income	Increases regional buying power	+	Lee, Florida, and Acs (2004)
	Increases opportunity costs for new venture creation	–	Love (1995)
Population density	Provides markets for starters congestion effects	+ –	Anyadike-Danes et al. (2005); Nerlinger (1998)
Proportion of resident aliens	Discrimination on first labor market	+	Lee et al. (2004)
	Lacking networks and resources	–	
Average firm size (incubator hypothesis)	Large firms spin out companies	+	Sutaria and Hicks (2004)
	Large firms erect entry barriers		
	Starters learn in smaller firms	–	Beesley and Hamilton (1984)

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

Table 3
Linear Regression Results, Summary Statistics of Local Parameter Estimates, and Results of the Monte-Carlo Test

	Linear Regression Results		Summary of Local GWR Parameters			Monte-Carlo Results		
Local Parameters	β_{lm}	SE_{lm}	Min.	Median	Max.	SD_{GWR}	p Value	$SD_{\mathrm{GWR}}/SE_{lm}$
Intercept	–1.573	3.523	–7.471	–0.097	21.260	9.337***	.000	2.650
Exit rate	0.448*	0.215	–0.408	–0.084	1.471	0.623***	.000	2.892
Unemployment rate	0.224*	0.102	–0.282	0.027	0.389	0.220***	.000	2.155
Average net income	0.841***	0.213	–0.111	0.546	1.347	0.382	.116	1.793
Population density	0.440	0.416	–0.085	0.790	1.606	0.360	.999	0.866
Proportion of resident aliens	–0.121	0.088	–0.345	–0.206	0.026	0.087	.997	0.989
Firm size	–1.092***	0.178	–1.678	–1.184	–0.799	0.221	.999	1.243
R^2	0.655		0.522	0.717	0.817			

Notes: GWR = geographically weighted regression; Max. = maximum; Min. = minimum.
Significance level: ***p < .001; *p < .05.

We illustrate the application of GWR as a tool to discover spatial heterogeneity with the example of an analysis of factors affecting start-up activity in Austria during 2006. As a response variable, the start-up activity is measured by the labor-market approach that standardizes the number of start-ups against the economically active population (Audretsch & Fritsch, 1994). As independent predictors, the exit rate (number of firm exits divided by number of existing firms), the average unemployment rate in 2005, the average net income in 2005, the average firm size in 2005, the population density at the beginning of 2006, and the proportion of resident aliens at the beginning of 2006 are considered in the model. Data at the regional level of 93 Austrian counties are provided by the Chamber of Commerce and Statistic Austria, and the analysis was performed with the free statistical environment R Version 2.8.0 (R Development Core Team, 2008). The GWR estimation function in R is provided by the software package spgwr Version 0.5-4 (Bivand & Yu, 2008).

First, we compare the results of the GWR with an approach that does not consider spatial heterogeneity. The comparison model is a global linear regression model. The global model yields an R^2 value of about 0.655, which represents an adequate fit. In the global linear model, the coefficients for exit rate (0.488), for unemployment rate (0.224), and for average net income (0.841) are significant and positive. The global model indicates that these variables have a positive relation with start-up activity over the whole study area, where for average firm size there is a significant negative relationship with start-up activity (−1.092). No significant effect could be identified for the population density or the proportion of resident aliens (see Table 3). However, the Moran I statistics (Doh & Hahn, 2008) show that the assumption of the linear model of an independent error term is violated. Therefore, we must consider spatial dependence in the analysis and must apply the spatial regression approach GWR.

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

Estimating the GWR Model

To investigate spatial heterogeneity, we followed the three-step framework discussed in the previous section. First, we estimated a GWR model. The adaptive bisquare kernel function (see Equation 3) was used to calculate the spatial weights. The adaptive bandwidth of the kernel was determined by AIC optimization that resulted in an optimal solution that considered 62.4% of the regions (about 58 of 93 counties) for each local regression model.

The summary statistics of the local parameters demonstrate that the estimates for most predictor variables change sign across the study area. Comparing the minima and maxima of local parameter estimates of the intercept, the exit rate, the unemployment rate, the average net income, the population density, and the proportion of resident aliens indicate that the direction of the relationship with the start-up rate (the response variable) varies in Austria. These findings indicate spatial heterogeneity (see Table 3).

To visualize how the local GWR parameters change over the study area, the GWR results can be plotted on a map (see Figure 1). A GIS system, such as the product family from Environmental Systems Research Institute Inc. or components in the statistical environment R (R Development Core Team, 2008) can be used to produce this kind of visualization. The level of local parameter estimates are indicated by the gradient of the lines, for example, the higher (lower) the parameter value, then the higher (lower) are the corresponding slopes. Negative local parameters are illustrated by a negative slope, whereas positive parameters are indicated by a positive slope. We used gray coloring to indicate when the local parameters are significant at the level of 5%, and white if they are not significant.

For example, the local parameters for the exit rate in Figure 1B range from –0.408 in eastern Austria to 1.471 in western Austria. Although there is a significant positive global effect, we detect a significant positive relationship in eastern Austria and a significant negative relationship with the start-up activity in the west. Similarly, we found sign changes for the impact of the unemployment rate (min. = –0.282, max. = 0.398), the average net income (min. = –0.111, max. = 1.347), the population density (min. = –0.085, max. = 1.606), and the proportion of resident aliens (min. = –0.345, max. = 0.026). In the cases of the average firm size, we did not detect sign changes but a regional variation of the strength of the relation with the response variable. The minima and maxima of local parameters for the average firm size are –1.678 and –0.799, respectively. The findings indicate that a neglect of spatial heterogeneity in a global model can lead to erroneous results due to counterbalancing or dilution effects (see Figures 1A–G and Table 3).

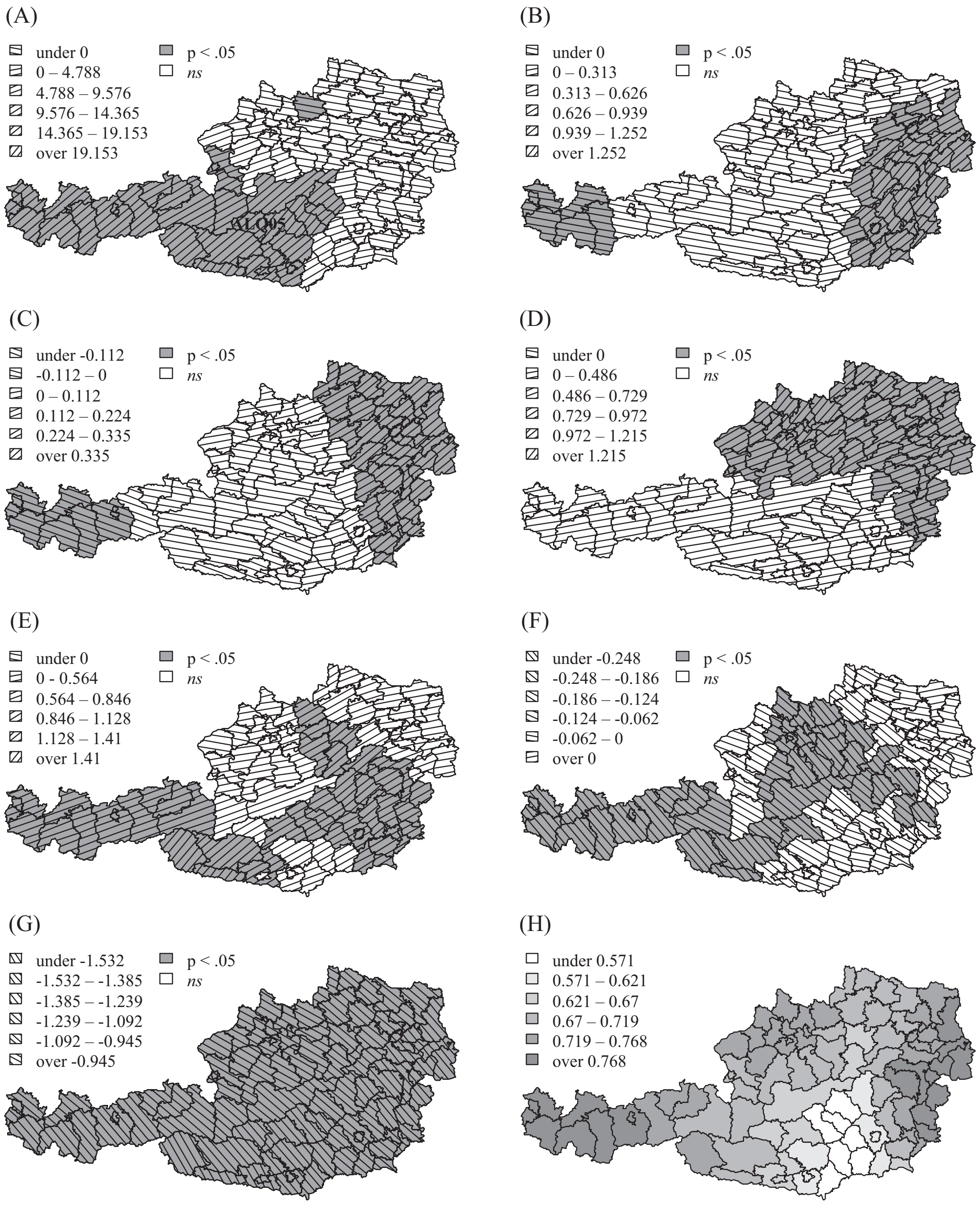
In addition to the local GWR parameters, we also plotted local goodness of fit statistics for each region i, the local R^2. This plot provides information about whether the fit is adequate over the whole study area. The local R^2 values are between 0.522 and 0.817. The fit measures show that the model has an adequate fit over the whole study area, with a better model fit in western, northern, and eastern Austria (see Figure 1H).

Determining Parameter Variability

Second, local parameter variability is assessed by calculating the standard deviations SD_k of local parameters. We compared the standard deviations with the standard errors of parameters in the global model to determine the degree of spatial variation following Brunsdon

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

Figure 1
Maps With Local Parameter Estimates of the Geographically Weighted Regression (GWR) Model. A, Intercept. B, Exit Rate. C, Unemployment Rate. D, Average Net Income. E, Population Density. F, Proportion of Resident Aliens. G, Average Firm Size. H, Local R^2.



Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

et al. (1998). The standard deviation of the exit rate with 0.623 is about 2.9 times the standard error of the global model with 0.215. Similarly, the standard deviation of unemployment rate (0.220) is about twice as high as the standard error of the global model (0.102). The standard deviation of the average net income and the average firm size also exceeds the standard error by a factor of 1.8, and 1.2, respectively. Only the standard deviations of population density (0.360) and proportion of resident aliens (0.087) are less than the standard errors of the global model with 0.416, and 0.088, respectively. Thus for four of seven variables and including the intercept, a high degree of variability can be reported. These findings are further indicators of spatial heterogeneity (see Table 3).

Testing Spatial Stability

Third, to test the significance of local parameter variation, a Monte-Carlo test (based on $L = 999$ permutations) was performed. To obtain independent test statistics, we ran the test procedure for each of the seven variables including the intercept separately. The standard deviations for the exit rate and the unemployment rate have the highest values in comparison to their resulting sampling distributions. Hence, the p values for these standard deviations equal 0.000, the differences in the local parameter estimates of these variables are significant, and the varying relationships across Austria have to be considered. For local parameters of average net income, population density, proportion of resident aliens, and average firm size, no significant spatial variability could be found (see Table 3).

Conclusion and Implications

Contributions of the Recognition of Spatial Heterogeneity to Entrepreneurship Research

In this article, we have discussed the phenomenon of *spatial heterogeneity* and illustrated that entrepreneurship research can benefit from analyzing the heterogeneous or homogenous nature of a study area by avoiding counterbalancing and dilution effects. In addition, entrepreneurship research can benefit by unambiguously establishing the geographical boundaries of generalizability. The identification of spatial heterogeneity is challenging, and only a few papers discuss this important concern. Methodologically, we have illustrated how spatial heterogeneity can be identified by using GWR. Finally, we suggest that researchers should become aware of potential spatial heterogeneity in their specific study areas.

By recognizing the importance of spatial heterogeneity, several contributions to entrepreneurship theory can be drawn from our example. First, when testing a theory in the context of a spatially heterogeneous study area, a researcher might erroneously not reject the null hypothesis (no effect), while in observable reality, nonrejection may be attributable to the *counterbalancing effects* of structural relationships in some local study areas. In our example, the overall effect of the proportion of resident aliens on start-up activity is not significant in the complete study area represented by the global model. However, the proportion of resident aliens is negatively significant in the western and central parts of Austria. If

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

a researcher wanted to test the impact of the proportion of resident aliens in the whole country, implications about the insignificance of the variable based on the results from the global model area would be misleading. By avoiding these kinds of misleading interpretations, considering spatial heterogeneity can enhance theory in the area of entrepreneurship.

Second, if different local study areas are characterized by a different strength of the relationship between predictor and response variables, *dilution effects* can result in a value of the overall model that does not represent the values of specific local study areas. The effect detected in a global model would be weaker or stronger than in some local study areas. In this example, the coefficient for the impact of average net income on new firm formation in the global model was 0.841 (significant at $p < 0.001$). However, Figure 1D indicates that significant relationships exist only in eastern Austria, while in western Austria, the relationship in the majority of study areas are weaker and nonsignificant (in a few cases even negative). In cases where entrepreneurship theory is concerned with the particular strength of association of predictor and response variables, the detection of varying strengths of relationships (even without a change in sign), within a broader study area, the recognition of spatial heterogeneity will contribute to the field.

Third, by *unambiguously identifying homogeneity*, stronger claims for generalization can be made. In the global model, the relationship between average firm size and start-up activity was significant, and GWR showed how this effect is present in all local study areas. Therefore, counterbalancing and dilution effects are not present and the results of the global model are valid in Austria as a whole.

Finally, the analysis of spatial heterogeneity is well suited for *exploratory research* (Fotheringham, 1992). If spatial heterogeneity is detected, questions about the mechanisms that cause the observed parameter variations can be asked. For example, why is it that the positive impact of the average net income is stronger in northeastern Austria than in southwestern Austria? Are cultural issues behind this finding, or is it due to a different industry structure? Fotheringham et al. (2002, p. 252) sum up the exploratory nature of GWR by stating, ‘‘GWR could be thought of, in fact, as a ‘spatial microscope’. The surfaces of parameter estimates from GWR raise new sets of questions about the general issue of parameter variations and the validity of global statements.’’

Additional Applications of Spatial Heterogeneity in Entrepreneurship Research

In our illustrative example, we showed how spatial heterogeneity could contribute to the analysis of new firm formation. In this example, the primary goal was to identify the boundaries of structurally similar local study areas. In the following example, we illustrate how spatial heterogeneity can be applied in the analysis of moderating relationships in the context of culture.

One area of entrepreneurship research that has commanded considerable attention is the investigation of the impact of culture on the entrepreneurial process (e.g., Garcia-Cabrera & Garcia-Soto, 2008). Culture, which can be defined as ‘‘a set of shared values, beliefs, and expected behaviours’’ (Hayton, George, & Zahra, 2002, p. 33) can play a role in analyses of the sources of country differences through aggregate measures of entrepreneurship,

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

characteristics of individual entrepreneurs, and on corporate entrepreneurship (Hayton et al., 2002).

In a model of the association between culture and entrepreneurship, Hayton et al. (2002) argue that cultural values can moderate the relationship between the institutional and economic context and entrepreneurship. For example, while a munificent economic environment is in principle supportive to entrepreneurship, individuals whose cultural values predispose them to act entrepreneurially will generate more entrepreneurial results than persons whose cultural values are not entrepreneurially predisposed. To analyze these moderating relationships, researchers could in principle identify culturally homogenous regions and use resulting regional dummy variables or indicators of cultural dimensions as moderator variables. Initially, it may seem easy to identify culturally homogenous regions because the idea of national cultures (Hofstede, 1991) implies that cultural differences are confined to national boundaries. However, cultural values can vary in a study area, not only in a national context but also in a regional context (Garcia-Cabrera & Garcia-Soto, 2008).

Cultural values exacerbate the *ex ante* delineation of study areas, particularly in the context of multinational cultures and multicultural nations. In addition, cultural values may not only be influenced by the cultural heritage of a group but also by institutional and cultural factors of groups that surround them (Tan, 2002). Moreover, even an ethnically homogenous group may differ in terms of motivation, demographics, and other variables (Cardon, Shinnar, Eisenman, & Rogoff, 2008), which could give rise to unexpected complex contingencies. Consequently, researchers who wish to analyze the impact of culture on the relationship between context factors and entrepreneurship could profit from GWR, a tool that considers the spatial variation of the relationships between response and predictor variables as well as their spatial dependence. GWR will identify structurally heterogeneous local study areas and researchers can assess whether cultural differences may be the source of potential variations.

Concluding Remarks

In addition to the analysis of new firm formation and culture, entrepreneurship researchers can benefit from the recognition of spatial heterogeneity in any analysis characterized by unclear boundaries of study areas and/or unexplored conditionalities that are spatially bound. Although the analysis of spatial heterogeneity is beneficial for entrepreneurship research, potential drawbacks of GWR should not be overlooked. Although in many data sets, spatial information is readily available, the need for georeferenced data makes primary data collection more costly. Software for GWR is available (e.g., GWR3x at http://ncg.nuim.ie/ncg/GWR/software.htm), however, diligent spatial analysis is still arduous. Additionally, particular regional values may be used for multiple local estimations in GWR. Therefore, multicollinearity may arise (Wheeler & Tiefelsdorf, 2005), which makes GWR more suitable for exploratory rather than confirmatory analyses.

The explicit consideration of spatial heterogeneity has the potential to contribute to theory creation and testing in entrepreneurship research as well as in other areas of the social sciences. An increased awareness of spatial heterogeneity will lead to exciting new research questions and results. We hope that this contribution sensitizes entrepreneurship scholars to

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

the tremendous opportunities that spatial econometrics presents in general (Doh & Hahn, 2008) and spatial heterogeneity, in particular, can bring to their research.

Note

1. A list of these studies together with the coding can be obtained from the authors.

References

Anselin, L. (1988). *Spatial econometrics. Methods and models.* Dordrecht: Kluver.

Anselin, L., & Florax, R. J. G. M. (1995). New directions in spatial econometrics—Introduction. In L. Anselin & R. J. G. M. Florax (Eds.), *New directions in spatial econometrics* (pp. 3-18). Berlin: Springer.

Anyadike-Danes, M., Hart, M., & O'Reilly, M. (2005). Watch that space! The County Hierarchy in firm births and deaths in the UK, 1980-1999. *Small Business Economics*, *25*, 273-292.

Audretsch, D. B., & Fritsch, M. (1994). On the measurement of entry rates. *Empirica*, *21*, 105-113.

Beesley, M. E., & Hamilton, R. T. (1984). Small firm's seedbed role and the concept of turbulence. *Journal of Industrial Economics*, *33*, 217-231.

Bivand, R., & Yu, D. (2008). spgwr: Geographically weighted regression (Version 0.5-4).

Brixy, U., & Grotz, R. (2007). Regional patterns and determinants of birth and survival of new firms in western Germany. *Entrepreneurship & Regional Development*, *19*, 293-312.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1998). Geographically weighted regression—Modeling spatial non-stationarity. *The Statistician*, *47*, 431-443.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, *39*, 497-524.

Carayannis, E. G., Assimakopoulos, D., & Kodo, M. (2008). *Innovation networks and knowledge clusters: Findings and insights from the US, EU, and Japan.* New York: MacMillan.

Cardon, M. S., Shinnar, R. S., Eisenman, M., & Rogoff, E. G. (2008). Segmenting the population of entrepreneurs: A cluster analysis study. *Journal of Developmental Entrepreneurship*, *13*, 293-314.

Davidsson, P. (2005). *Researching entrepreneurship.* New York: Springer.

Debarsy, N., & Dejardin, M. (2008). *What do we know about spatial entry?* Jena Economics Research Papers 2008-29, Jena, Germany.

Doh, J. P., & Hahn, E. D. (2008). Using spatial methods in strategy research. *Organizational Research Methods*, *11*, 659-681.

Eurostat. (2009). Nomenclature of territorial units for statistics—NUTS statistical regions of Europe. Retrieved January 15, 2009, from http://ec.europa.eu/eurostat/ramon/nuts/home_regions_en.html

Fotheringham, A. S. (1992). Exploratory spatial data analysis and GIS. *Environment and Planning*, *4*, 1675-1678.

Fotheringham, S. A., Brunsdon, C., & Charlton, M. (2000). *Quantitative geography: Perspectives on spatial data analysis.* London: Sage Publications.

Fotheringham, S. A., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships.* Chichester: Wiley.

Fritsch, M. (2004). Entrepreneurship, entry and performance of new business compared in two growth regimes: East and West Germany. *Journal of Evolutionary Economics*, *14*, 525-542.

Garcia-Cabrera, A. M., & Garcia-Soto, G. (2008). Cultural differences and entrepreneurial behaviour: An intra-country and cross-cultural Analysis in Cape Verde. *Entrepreneurship and Regional Development*, *20*, 451-483.

Hayton, J. C., George, G., & Zahra, S. A. (2002). National culture and entrepreneurship: A review of behavioral research. *Entrepreneurship: Theory and Practice*, *26*, 33-52.

Hofstede, G. (1991). *Cultures and organizations.* New York: McGraw-Hill.

Julien, P. A. (2007). *A theory of local entrepreneurship in the knowledge economy.* Cheltenham: Edward Elgar.

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010

Karlsson, C., & Dahlberg, R. (2003). Entrepreneurship, firm growth and regional development in the new economic geography: Introduction. *Small Business Economics*, *21*, 73-76.

Lee, S. Y., Florida, R., & Acs, Z. J. (2004). Creativity and entrepreneurship: A regional analysis of new firm formation. *Regional Studies*, *38*, 879-891.

Love, J. H. (1995). The measurement of entry rates: Reconsideration and resolution. *Empirica*, *22*, 151-157.

Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. *Academy of Management Review*, *21*, 135-172.

Miller, D. (1983). The correlates of entrepreneurship in three types of firms. *Management Science*, *29*, 770-791.

Nerlinger, E. A. (1998). *Standorte und Entwicklung junger innovativer Unternehmen: Empirische Ergebnisse fuer West-Deutschland.* Baden-Baden: Nomos-Verlagsgesellschaft.

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reynolds, P., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies*, *28*, 443-456.

Ritsilä, J, & Tervo, H. (2002). Effects of unemployment on new firm formation: Micro-level panel data evidence from Finland. *Small Business Economics*, *19*, 31-40.

Rogers, E. M. (1983). *Diffusion of innovations.* New York: Free Press.

Sutaria, V., & Hicks, D. A. (2004). New firm formation: Dynamics and determinants. *Annals of Regional Science*, *38*, 241-262.

Tan, J. (2002). Culture, nation, and entrepreneurial strategic orientations: Implications for an emerging economy. *Entrepreneurship: Theory and Practice*, *26*, 95-111.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*, 234-240.

Tödling, F., & Wanzenböck, H. (2003). Regional differences in structural characteristics of start-ups. *Entrepreneurship & Regional Development*, *15*, 351-370.

van Oort, F. G., & Atzema, O. A. L. C. (2004). On the conceptualization of agglomeration economies: The case of new firm formation in the Dutch ICT sector. *The Annals of Regional Science*, *38*, 263-290.

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographic Systems*, *7*, 161-187.

Robert J. Breitenecker is assistant professor of entrepreneurship and innovation management at Klagenfurt University. His research interests include multivariate and spatial statistical methods, entrepreneurship, and innovation management. He has published in *International Journal of Business and Globalisation* and *Zeitschrift für Betriebswirtschaft*.

Rainer Harms is assistant professor for entrepreneurship at NIKOS, University of Twente (The Netherlands). He has published in journals such as *Entrepreneurship and Regional Development, European Management Journal*, and *Zeitschrift für Betriebswirtschaft*. He is currently working on development paths of new ventures based on the configuration approach.

Downloaded from orm.sagepub.com at Universiteit Twente on October 4, 2010