

This is the version of the article accepted for publication in Small Business Economics published by Springer: <https://doi.org/10.1007/s11187-014-9613-1>
Accepted version downloaded from SOAS Research Online: <http://eprints.soas.ac.uk/19129>

Entrepreneurship across Time and Space: Empirical Evidence from Korea

Eunsuk Hong*

Department of Financial and Management Studies
SOAS, University of London,
Thornhaugh Street, Russell Square, London WC1H 0XG, UK
Tel. +44 20 7898 4564
Fax +44 28 7898 4089
Email: e.hong@soas.ac.uk

In Hyeock Lee

Department of Management, Quinlan School of Business
Loyola University Chicago
820 N. Michigan Ave., Chicago, IL 60611, USA
Tel. (312) 915 7656
Fax (312) 915 6988
E-mail: ilee1@luc.edu

Laixiang Sun

Department of Geographical Sciences
University of Maryland
2181 Samuel J. LeFrak Hall, College Park, MD 20742, USA
Tel. (301) 405 8131
Email: lsun123@umd.edu

Department of Financial and Management Studies
SOAS, University of London
Thornhaugh Street, Russell Square, London WC1H 0XG, UK

&

Richard Harrison
Business School
University of Edinburgh
29 Buccleuch Place, Edinburgh, EH8 9JS, UK
Tel. +44 131 651 5549
Email: R.Harrison@ed.ac.uk

* Corresponding author

Entrepreneurship across Time and Space: Empirical Evidence from Korea

Abstract

This paper investigates the temporal and spatial dynamics of business start-up activities and their determinants. It integrates three perspectives in explaining regional variations of start-ups: (i) spatial heterogeneity that characterizes regional differences in promoting or conditioning start-up activities, (ii) temporal dependence that features a self-augmenting and self-reinforcing effect of start-up activities, and (iii) spatial dependence that portrays inter-regional interaction of start-up activities across proximate regions. A spatial dynamic panel modeling analysis of the determinants of new manufacturing ventures created in sub-national regions of South Korea confirms that, in addition to the importance of regional characteristics, both temporal and spatial dependences of start-up activities are simultaneously in force and play statistically significant roles. To address the joint endogeneity issue of temporal and spatial dependences, we employ the system GMM estimator, which leads to much improved explanation of inter-regional variations in firm creation activities.

Keywords: Spatial dependence, Temporal Dynamics, New firm creation, Spatial dynamic panel, Economic Geography, Korea

JEL Classification Codes: M13, L26, C18

1 Introduction

Despite the claim of the ‘death of distance’ in a globalizing world (Cairncross 1997), interest in the geography of economic activities has increased in recent years. This growing interest manifests itself in such topics as the co-location of firms and the relationship it has with economic growth (Krugman 1991), the clustering of economic activities (Porter 1998), and the role of geography in the strategic management of firms (Baum and Sorenson 2003). One of the emphases in this stream of the literature is on the regional variations of entrepreneurial activity and their determinants (e.g., Reynolds et al. 1995; Armington and Acs 2002; Fritsch and Falck 2007; Lasch et al. 2013). This is not surprising because on the one hand, the location of both new and already existing firms is essential for the regional economic development (Audretsch et al. 2006; Fritsch and Mueller 2008) and on the other hand, differences in the rates of entrepreneurship and in their determinants have been pronounced (Acs et al. 2007; Brixy and Grotz 2007). Despite extensive studies in the regional variations of firm start-ups and their determinants, the existing analyses are mainly static. Recently, there has emerged a small set of empirical publications which deal with the temporal or spatial dynamics of start-up activities. However, this literature has developed into two separate tracks of temporal versus spatial dependence and has paid little attention to the joint effect of these two dynamics. Given the established recognition that entrepreneurship is a process of collective, network-based activities (Steyaert and Katz 2004; Schienstock 2007; Trettin and Welter 2011; Fredin forthcoming), it is surprising that the joint temporal-spatial context of entrepreneurship is still missing in the debate.

This research aims to fill this niche and to address the joint temporal and spatial dependences in the entrepreneurship research. We integrate the three existing explanatory tracks in the literature: (1) spatial heterogeneity, (2) temporal dependence, and (3) spatial

dependence. First, the spatial heterogeneity perspective argues that regional differences in start-up activities can be explained by the presence of substantially different regional characteristics which condition the formation of entrepreneurship capital in the region (Audretsch and Keilbach 2004; Pijnenburg and Kholodilin 2014), such as income level, labor cost, land cost, market size, density of related firms, infrastructure, and other endowments conducive to the creation of new businesses. Since most of these factors are spatially sticky and relatively time-inert, the regional differences in the intensity of entrepreneurial activities tend to be persistent over time.¹ This perspective is in line with economics of agglomeration, in which spatial heterogeneity is essential for explaining a non-uniform distribution of economic activities across locations and for the specializations of certain regions and cities (Arthur 1994; Krugman 1991; Fujita and Thisse 2002).

Second, the temporal dependence perspective suggests path dependence in start-up activities themselves, meaning that the level of start-up activities in the previous period(s) may have prediction power on its observation in the current period (e.g., Zucker et al. 1998; Fritsch and Mueller 2007; Andersson and Koster 2011). This perspective is in line with evolutionary economic geography, in which the current state of affairs has emerged from and has been constrained by previous states of affairs (Boschma and Frenken 2006). The causes for path dependence in start-ups found in the literature are spatially bounded externalities and local institutions (Martin and Sunley 2006) and also cumulative causation that the pursuit of entrepreneurial opportunities feeds further opportunities (Holcombe 2007).

Third, the spatial dependence perspective highlights the spatial interdependency in start-up activities and argues that new firm creation activities in close proximity have prediction power on similar activities in the focal region (Plummer 2010). This

¹ See Stam and Lambooy (2012) for a review of literature on the spatial heterogeneity perspective.

perspective is motivated by migration, interregional trade, technology and knowledge spillovers, business clusters spreading across neighboring regions, diffusion of business norms as well as social and institutional networks (Audretsch et al. 2006; Pijnenburg and Kholodilin 2014). Despite its importance, the issue of spatial dependence in entrepreneurial activities has not been sufficiently examined. For example, according to Breitenneckar and Harms (2010), of the 40 publications they surveyed, Oort and Atzema (2004) is the (only) one with an attempt to take into account spatial dependence of new firm formation of Dutch high-technology enterprises in parameters estimation by adopting a spatial lag model. Recently, Plummer (2010) presents a systematic illustration of spatial modeling techniques for entrepreneurship research.

From a methodological perspective, if the level of current start-up activities in a focal region can be predicted not only by the past level of such activities in the region (i.e., temporal dependence) but also by the levels of start-up activities in the neighboring regions (i.e., spatial dependence), while controlling for other local determinants (i.e., spatial heterogeneity), then the corresponding measures are serially and spatially autocorrelated. This double autocorrelation invalidates the classical statistical methods such as ordinary least square and panel regressions with fixed or random effects because these classical methods build on the premise of random sampling that the observation of the variables are (serially and spatially) independent (Baltagi 2008). To address the joint temporal and spatial dependences of start-up activities, we specify a spatial dynamic panel model and employ the system Generalized Method of Moment (GMM) estimator to estimate the model. The system GMM estimator is arguably the best available estimator which is capable of addressing the joint endogeneity issue of temporal and spatial dependences.

Using a unique population dataset of 44,434 newly created small manufacturing firms in 234 regions of South Korea (hereafter Korea) during 2000-2004, we find that the regional activity level of new firm creation in Korea is indeed serially and spatially correlated over time and across locations. This finding further provides statistical justification for considering the joint endogeneity issue of the serially and spatially lagged dependent variable in the empirical estimation. A comparison of the system GMM estimation results with those from alternative estimators such as the pooling regression with OLS and spatial panel regression with maximum likelihood (ML) indicates that by addressing the joint endogeneity of temporal and spatial dependences, the system GMM estimator produces a much improved explanation of inter-regional variations in the level of new firm creation activities.

In addition to the data availability, the following two reasons further strengthen the suitability of Korea for this study. First, after the financial crisis in the late 1990s, the Korean government initiated the provision of strong incentive packages to promote prospective entrepreneurs' founding activities of new ventures. This initiative led to a boom in start-up activities, which provided excellent opportunities for researchers to explore potential interconnectedness of new firm creation across time and space. Second, Korea has been recognized as one of the most entrepreneurial societies (Reynolds et al. 2002) with the third highest level of entrepreneurial activity (Bosma et al. 2009), and is among the top 20 most competitive economies (Schwab 2009) in the world. However, much is yet to be known on temporal and spatial patterns of the start-ups of Korean entrepreneurial firms, since they have been relatively underrepresented in the entrepreneurship literature.

The rest of the paper is organized as follows. Section 2 first summarizes the theoretical foundations of spatial heterogeneity, and the temporal and spatial dependences

in entrepreneurship. It then defines the key variable of the research – the level of domestic entrepreneurial activities, and statistically tests the serial and spatial dependences of this variable. Section 3 develops an integrated spatial dynamic panel model for the determinants of the regional variations of start-ups, which incorporates the perspectives of temporal dynamics and spatial dependence, whilst controlling for the heterogeneity of regional characteristics. Section 3 also specifies the major independent variables and suggests the estimation method. Section 4 reports and interprets the results of the estimations. Finally, Section 5 discusses the theoretical and policy implications of the empirical findings and presents the concluding remarks.

2 Spatial heterogeneity, the temporal and spatial dependences in entrepreneurship

2.1 Theories

Spatial heterogeneity. The spatial heterogeneity perspective of entrepreneurial activities puts an emphasis on the socio-spatial context of entrepreneurship (Steyaert and Katz 2004; Trettin and Welter 2011). In other words, it veers from the big-man theory of entrepreneurship towards a collective action and network-based understanding of entrepreneurship (Fredin forthcoming). One of the intuitive reasons is that innovations are rarely the achievement of one individual but rather of a group of individuals (Graf 2011). In terms of knowledge flows, although the importance of interregional networks is undeniable, most contacts are local, especially for innovative entrepreneurial activities where knowledge flows between talented individuals are crucial (O'Donnell et al. 2001; Nijkamp 2003; Smith et al. 2005). In addition, the theoretical and empirical literature of economic geography and spatial economics has provided strong reasons and evidence for the uneven distribution of production factors across regions, which condition the uneven

development of physical and human capitals across regions. For example, in urban economics, spatial heterogeneity acts as a driver for inducing a structural instability of economic activities across space and the specializations of certain regions and cities (Arthur 1994; Krugman 1991; Fujita and Thisse 2002). In this regard, a central issue that the entrepreneurship literature attempts to explore is not only the phenomena of start-up activity but also how the relationships conditioning start-up activities differ across regions (Breiteneckar and Harms 2010).

Temporal dependence. In evolutionary economic geography, the notion of ‘time’ is conceptualized as a path dependent process, in which ‘the current state of affairs cannot be derived from current conditions only, since the current state of affairs has emerged from and has been constrained by previous states of affairs’ (Boschma and Frenken 2006, p. 180). In the literature explaining technological adaptation processes and industry evolution, temporal or path dependence is attributed to the quasi-irreversibility of investments, economies of scale, and technical interrelatedness or the need for system compatibility (David 1985; Arthur 1994). A manifest of this notion in the regional entrepreneurship literature suggests a self-augmenting process (Fornahl 2003) and a self-reinforcing effect of entrepreneurship (Minniti 2005), indicating that current entrepreneurial activities are partly a response to those in previous periods (Andersson and Koster 2011). According to this self-augmenting and self-reinforcing approach to entrepreneurship, regional entrepreneurial history plays an important role in determining the level of entrepreneurial activities via demonstrating, coaching and learning (Shane 2003; Arenius and Minniti 2005).

Spatial dependence. Spatial dependence generally refers to ‘the existence of a functional relationship between what happens at one point in space and what happens elsewhere’ (Anselin 1988, p.11). In the new economic geography developed by Krugman

(1991), the notion of ‘space’ has been an important factor that fosters a spatial dependent process of activities. For example, the role of agglomerative spillovers in the urban economics literature has emphasized the benefits that firms obtain from their neighbors over space (Brueckner 2011). According to the perspective on the spatial dimension of entrepreneurship capital, the accessibility to entrepreneurship capital and competition induced by entrepreneurial activities are decisive to firm creation activities and they are not bounded by the borders of administrative jurisdictions (Pijnenburg and Kholodilin 2014). While innovative activity tends to cluster spatially, meaning that regions with large amounts of innovative activity typically experience a high level of knowledge-based entrepreneurial activity (Audretsch and Feldman, 1996), Porter (1998) argues that clusters may cross state or even national boundaries. Therefore, the decision on where to locate an entrepreneurial new firm is not only based on the entrepreneurship capital of a given region, but also on the spatial feature of the cluster which can be spread across several regions.

2.2 Preliminary statistical evidence in Korea

The key variable for this research is the level of domestic entrepreneurial activities in region i and year t . This variable will be the dependent variable in the modeling sections of the research and is measured as the logarithm of one plus the total number of newly created entrepreneurial firms in the manufacturing sector in region i and year t , denoted as $\ln Firm_{it}$.² The data are obtained from the Factory Establishment and Management Information System (FEMIS) database of manufacturing industries (<http://www.femis.go.kr>), compiled by the Korean Ministry of Trade, Industry and Energy

² The adoption of adding one to the count variable, which keeps observations with zero number of new firms after taking logarithm, is popular in the literature (e.g., Head et al. 1995; Maitland et al. 2005).

(<http://motie.go.kr>). The database provides a complete and reliable population of newly created manufacturing firms in Korea. Over the period of 2000-2004, the FEMIS database provide registration data for a total number of 58,564 new firms.³ Following the existing entrepreneurship literature, we take those new firms with number of employees equal to or less than 200 as entrepreneurial firms (Brush and Vanderwerf 1992; St-Jean and Audet 2012) and consequently, 44,434 start-ups remain in the final sample. Because the unit of analysis is defined at the regional level and Korea has 234 administrative county- and city-level regions, the variable $\ln Firm_{it}$ has a total number of 1,170 region-year observations (234 sub-national regions \times 5 years).

Insert Figures 1 and 2 and Table 1 about here

The scatter plots in Figure 1 show the visualized evidence of high and statistically significant autocorrelation of the entrepreneurial activities between year t and $t - 1$. The four correlation coefficients range between 0.851 and 0.935 and are all statistically significant at the 1 % level. In addition, the Wooldridge test statistics for autocorrelation in panel data gives a result of $\chi^2 = 97.113$ and $P \leq 0.0001$, indicating the presence of path dependence in new firm start-up activities in Korea.

To show spatial dependence in entrepreneurship activities, we employ the Moran scatter plots (Figure 2) and Moran's I statistics for spatial autocorrelation (Table 1). The

³ We chose 2000-2004 for the following reasons. First, Korea suffered from a severe financial crisis between late 1997 and early 1999. Economic activities before 2000, including the creation and location decisions of new ventures, would have been distorted due to the crisis. Second, the government amended the Korean Standard Industrial Classification (Korean SIC) Code in January 2000. To ensure the consistency of the types of industries to be included in our sample, we opt to use the data from 2000 onwards. Third, 2004 is the most recent year of the data to which we have access.

Moran scatter plot is a tool for visual exploration of spatial autocorrelation between the original variable and the spatially weighted average of the variable across neighboring locations (Anselin 1996). Moran's I -statistics, which is equivalent to the slopes of the best-fit-lines in the Moran scatter plots, is defined as

$$I = \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$

in which x_i and x_j are the observations for region i and j of the variable under analysis (i.e., $\ln Firm_{it}$ in this study); \bar{x} is the average of the variable across regions; and w_{ij} is the i - j element of the row-standardized \mathbf{W} matrix of weights. We construct the weight matrix \mathbf{W} based on the physical contiguity between regions, leading to a binary and symmetric matrix in which the cell (i, j) has value 1 if regions i and j are next-door neighbors, and value zero otherwise (i.e., first-order contiguity matrix). As both Figure 2 and Table 1 clearly indicate, the slopes of the best-fit-lines in the Moran scatter plots are positive and the values of z -statistics ensure that the Moran I statistics are statistically significant at the 1% level.

The above results provide statistical evidence that in the case of Korea, the current level of new firm start-up activities in a focal region is correlated not only with the past level of such activities in the region, but also with the levels of start-up activities in their neighboring regions. In the following section, we propose an integrated empirical model for firm creation activities that are subject to both temporal and spatial dependences.

3 Model specification and estimation methods

3.1 Model specification

The benchmark setting of our empirical model is the standard spatial heterogeneity model focusing on regional determinants of entrepreneurial activities, which is arguably the most widely adopted setting in the literature (Breiteneckar and Harms, 2010). The benchmark model can be presented as follows:

$$\ln Firm_{it} = \beta_0 + \beta_1 \ln \mathbf{LocationVariables} + \eta_i + \mu_t + \varepsilon_{it} \quad (1)$$

in which $\ln Firm_{it}$ is defined as before (Section 2.2); the vector '*Location Variables*' captures the standard regional determinants of start-ups; and the disturbance terms consist of the unobserved regional fixed effect that is constant over time (η_i), the unobserved time effect that is common for all regions (μ_t), and the transitory errors (ε_{it}) that may vary across regions and over time with a zero mean value. All location variables are transformed in natural log as typically adopted in the literature.

In recognition of the strong path dependence in entrepreneurial activities (Section 2), we incorporate a temporal dynamic dimension into Equation (1). This leads to

$$\ln Firm_{it} = \beta_0 + \beta_1 \ln \mathbf{LocationVariables} + \gamma \cdot \ln Firm_{i,t-1} + \eta_i + \mu_t + \varepsilon_{it} \quad (2)$$

in which $\ln Firm_{i,t-1}$ is the standard first-order temporal autoregressive term and the coefficient γ reflects the strength and sign of the temporal autoregressive term.

The incorporation of the spatial dependence of entrepreneurial activities into Equation (2) leads to the full model of this research as follows:

$$\begin{aligned} \ln Firm_{it} = \beta_0 + \beta_1 \ln \mathbf{LocationVariables} + \gamma \cdot \ln Firm_{i,t-1} \\ + \rho \cdot \mathbf{W} \cdot \ln Firm_{it} + \eta_i + \mu_t + \varepsilon_{it} \end{aligned} \quad (3)$$

where $\mathbf{W} \cdot \ln Firm_{it}$ captures a contiguity-weighted average of the number of new firms created in all other $j \neq i$ regions (i.e., entrepreneurial activities taking place in neighboring regions); and ρ is a parameter to be estimated that indicates the strength and

sign of the spatial autoregressive term, $\mathbf{W} \cdot \ln Firm_{it}$. As presented in Section 2.2, \mathbf{W} denotes a first-order contiguity weight matrix and $\mathbf{W} \cdot \ln Firm_{it}$ refers to a contiguity-weighted average of entrepreneurial activities in all other $j \neq i$ regions that are directly bordering the focal region i .

3.2 Location variables

'Location Variables' represent the key regional characteristics. In line with the literature on the regional determinants of start-up activity, we focus on four aspects of such locational characteristics: regional market size, the degree of agglomeration, local wage level, and the size of land available for building factories. In addition, we also introduce industrial complex dummies and yearly dummies to control for the different industrial policies initiated by local governments and year-specific characteristics, respectively. The data for the four major locational variables and industrial complex dummies are from governmental statistics on regional economies of the Korean National Statistics Office (<http://kosis.nso.go.kr/>). We recognize that it takes time for prospective entrepreneurs to collect and assess information, to make the decision and to materially establish a new firm. This consideration justifies the use of one-year lagged for 'Location Variables' in the model specification (e.g., Johnson and Parker 1996; Fritsch and Falck 2007). Therefore, the final specification of the empirical model is as follows:

$$\begin{aligned} \ln Firm_{it} = & \beta_0 + \beta_1 \ln GRP_{i,t-1} + \beta_2 \ln Agglomeration_{i,t-1} + \beta_3 \ln Wage_{i,t-1} \\ & + \beta_4 \ln Land_{i,t-1} + \gamma \ln Firm_{i,t-1} + \rho \cdot \mathbf{W} \cdot \ln Firm_{it} + \eta_i + \mu_t + \varepsilon_{it} \end{aligned} \quad (4)$$

In Equation (4), $\ln GRP_{i,t-1}$ captures the local market size and GRP stands for the total amount of gross regional product (in million KRW, US\$1 = 1,130 KRW) from manufacturing industries in each region. We expect $\ln GRP_{i,t-1}$ to have a positive effect on the level of new firm creation activities (Keeble and Walker 1994).

$\ln Agglomeration_{i,t-1}$ is measured as the total number of firms operating in each region across all manufacturing and service sectors, representing the degree of concentration or clustering of firms (Stuart and Sorenson 2003; Lasch et al. 2013). We expect a positive impact of agglomeration economies on new firm creation activities due to the availability of spatial and inter-industry linkages, accumulation of production factors, availability of information, and knowledge/technology spillovers. On the other hand, it is worth noting that although the majority of studies found a positive effect of agglomeration on entrepreneurial activities, a negative congestion effect (e.g., severe competitions among local firms for accessing financial resources and network partners, hiring workers, and sharing common infrastructure) is possible and it might outweigh the positive agglomeration advantages (Folta et al. 2006).

$\ln Wage_{i,t-1}$ captures the level of local labor cost and $Wage$ is measured as the average monthly wage per employee (in million KRW) of each region. A negative relationship between wage and the level of entrepreneurial activities would exist because rises in wage rates increases the opportunity costs of self-employment and also the cost of hiring workers (Fritsch and Falck 2007), and more particularly, the competitiveness of the labor-intensive manufacturing sector highly depends on a cheap labor force. Nevertheless, Zucker et al. (1998) shows a positive relationship between local wage level and the level of new firm creation activities in US biotechnology sectors. This is because a high wage level in a region may indicate a high level of human capital which is fundamental for the biotechnological sectors.

It is worth highlighting that for the start-ups of new manufacturing firms, the usual regional size issue is not about the total territory of the region, but land areas available for building factories in the region. Therefore we employ $\ln Land_{i,t-1}$, which is measured as

the total land area (in km²) available for building factories in region i and year $t - 1$, to capture this size effect and also the effect of zoning policies. We expect $\ln Land_{i,t-1}$ to have a positive effect on entrepreneurial activities because an increase in land availability for building factories may lead to an increase in potential locational sites for accommodating new manufacturing firms in the region. Table 2 summarizes the descriptive statistics and presents a correlation matrix of the variables.

We note the mixed results on testing the relationship between unemployment and new firm creation activities in the literature (Reynolds et al. 1995; Johnson and Parker 1996; Armington and Acs 2002; Fritsch and Falck 2007; Fritsch and Mueller 2007). This ambiguity may result from the countervailing influences of unemployment on new firm creation activities in a given region, e.g., on the one hand, high unemployment increases new firm start-up activities due to the low opportunity cost from the lack of other employment opportunities. On the other hand, a high level of employment might indicate low market demand, thus negative prospects for new business opportunities. We would also like to report that the variation in unemployment rates across Korean 234 regions over the period of 1999-2003 was surprisingly low and that the introduction of unemployment rates as an additional locational variable does not add meaningful explanatory power.

Insert Table 2 about here

3.3 Estimation methods

Three estimators are necessary for this research. The first one is the pooling regression with ordinary least squares (OLS), which serves as a benchmark regression for running standard spatial diagnostics tests (Anselin et al. 1996). These spatial diagnostics tests will provide justification for employing the second estimator: spatial autoregressive

regression with maximum likelihood (ML) to address the endogeneity issue of the spatially lagged dependent variable. Finally, the inability of spatial lag ML estimator to address the endogeneity issue of serially lagged dependent variable leads to the application of the third estimator: a combined spatial and dynamic panel regression with system GMM.

As has become well-known in modern econometrics, the pooling OLS estimation of the coefficient ($\hat{\gamma}$) on a serially lagged dependent variable ($\ln Firm_{i,t-1}$) is likely to produce inconsistent and upward-biased results owing to the positive correlation between $\ln Firm_{i,t-1}$ and η_i (Hsiao 2003). The inclusion of a spatially lagged dependent variable ($\mathbf{W} \cdot \ln Firm_{i,t}$) on the right-hand side of the model further causes simultaneity and endogeneity problems, which would make OLS estimators biased and inconsistent (Anselin 1988). Therefore, both the serially and spatially lagged dependent variables need to be treated as endogenous and the proper estimation of their coefficients ($\hat{\gamma}$ and $\hat{\rho}$) should account for this endogeneity issue in an explicit manner. The spatial autoregressive regression with maximum likelihood (ML) estimator can effectively address the endogeneity issue of the spatially lagged dependent variable; nevertheless, the existing spatial ML estimators are not designed to address the endogeneity problem caused by the inclusion of the serially lagged dependent variable (Elhorst 2003).

The difference GMM estimator proposed in Arellano and Bond (1991) removes the fixed effects by first differencing the data then instrumenting all potentially endogenous differences with their past levels. However, Blundell and Bond (1998) shows that if the series are persistent, difference GMM performs poorly because past levels convey little information about future changes. As a result, lagged levels are weak instruments for the first differenced variables. To increase efficiency, Blundell and Bond (1998) develop the

system GMM which adds moment conditions and combines ‘the original level equation’ and ‘the difference equation’ as a system of equations, with first-differences instrumented on lagged levels and with levels instrumented on first differences. By doing so, the system GMM makes instruments exogenous to the fixed effects and increases efficiency.

Kukenova and Monteiro (2008) compare the performance of several econometric estimators for panel models with both spatial and temporal dependences, including the spatial ML (Anselin 1988), the spatial dynamic ML (Elhorst 2005), the spatial dynamic quasi-ML (Yu et al. 2008), and the system GMM (Blundell and Bond 1998). Their Monte-Carlo investigations suggest that an application of the system GMM to spatial dynamic panel models, e.g., Equation (4) in our case, appears to be the best estimator as it deals with the joint endogeneity problem of serial and spatial dependences and corrects for the potential endogeneity of other explanatory variables. A recent study by Sun et al. (2010) on cross-province growth in China over the period of 1980-2005 also confirms the advantage of the system GMM estimator in comparison with other alternatives.

In order to guarantee the validity and relevance of the selected IVs and the consistency of the system GMM estimator, three validity tests are employed: first, the overall validity of the instruments are tested by the Hansen’s J test of over-identifying restrictions; second, following Roodman (2009), the Difference-in-Hansen tests for the full set of instruments for the level equation as well as for the subset based on the dependent variable is conducted, with the number of instruments reported; and third, the first-order and second-order serial correlation in the first-differenced residuals are tested. If the original error terms are not serially correlated, the first-order serial correlation test should be significant and the second-order one should be insignificant. In addition to the validity tests, a finite-sample correction for the two-step covariance matrix proposed by Windmeijer (2005) is applied.

4. Empirical results

Table 3 reports the results of the three estimators discussed above. The pooling OLS regression facilitates several spatial diagnostics tests for selecting relevant spatial econometric models that model the corresponding spatial dependence explicitly (Anselin et al. 1996). There are two primary types of spatial dependence (Anselin 1988; Plummer 2010). First, substantive spatial dependence is attributed to spatial interaction effects of social and economic phenomena (i.e., the entrepreneurial activity observed in one location depends on those observed in neighboring locations). Second, apparent spatial dependence is caused by measurement errors or noise in the data (i.e., the error terms across different spatial units are correlated). If the dependent variable is spatially correlated, then a spatial lag model in which the spatially lagged dependent variable is added to the right-hand-side of the linear equation can be used. An alternative way to analyze the spatial dependence is called a spatial error model in which the spatial dependence is assumed to be present in the errors of the linear model (Anselin 1988). Given the results of the Moran's I analysis above, it is no surprise that both LM lag and Robust LM lag statistics are significant at the 1% level and thus clearly indicate the existence of substantive spatial dependence in the level of new firm creation activities. This finding provides a strong justification for adopting the spatial lag model as the proper specification for estimating Equation (4). The statistical insignificance of the LM error and Robust LM error statistics indicate that apparent spatial dependence in the errors terms is not an issue in modeling the level of firm start-ups and therefore, the spatial error model is not an appropriate choice for the model specification.

In the spatial lag ML estimation, the statistically significant results of the Wald, LR, and LM tests further confirm the spatial autocorrelation of the dependent variable. The

estimate of the spatial lag parameter ($\hat{\rho}$) is positive (0.039) and statistically significant at the 10% level. The result provides evidence for the inter-regional spillovers between start-up activities in neighboring regions and those in the focal region (Plummer 2010).

For other variables, both the pooling OLS and spatial lag ML produce statistically equivalent results. With regard to the four locational variables, the coefficients on market size ($\ln GRP_{i,t-1}$) and the level of agglomeration ($\ln Agglomeration_{i,t-1}$) are positive and significant at the 1% and 10% levels, respectively, in line with the expectation of the research. However, the coefficients on the local wage rate ($\ln Wage_{i,t-1}$) and land available for building factories ($\ln Land_{i,t-1}$) are statistically insignificant, which are not in line with the expectation of this research. With regard to the serially lagged dependent variable, its coefficient in both regressions is significantly positive, suggesting the existence of strong path dependence in start-up activities themselves. This is in line with previous results for Germany (Fritsch and Mueller 2007) and Sweden (Andersson and Koster 2011). However, the estimated values of $\hat{\gamma}$ in both regressions are equally high, at 0.763 in the pooling OLS and 0.761 in spatial lag ML. This calls for caution because we know that the pooling OLS estimation tends to produce inconsistent and upward-biased results for the autoregressive term (Hsiao 2003). This caution implies that, in the presence of both serially and spatially lagged dependent variables, a control for the endogeneity of spatial dependence alone may not be sufficient for correcting the upward bias on $\hat{\gamma}$.

The last column of Table 3 reports the results of system GMM estimation. The column shows that the system GMM estimation passes all the specification tests of Hansen's J , Difference-in-Hansen, the first-order and second-order serial correlation tests, indicating that the IVs selected in the regression are statistically valid and the original

error terms are not serially correlated. As a result, the system GMM is qualified to be the preferred estimator in comparison with the other two.

In comparison with the pooling OLS and Spatial lag ML, the improvement brought in by the system GMM estimation is multifold. First, it reduces the value of the coefficient on serially lagged dependent variable ($\hat{\gamma}$) by about one third, from more than 0.760 to 0.414, implying that a 10% increase in the number of new firms created in year $t-1$ is now associated with a 4.14% increase in the number of new firms created in year t . This is an indication of a significant correction for the upward bias on $\hat{\gamma}$ produced by the pooling OLS and spatial lag ML. Second, the system GMM estimation raises the value of $\hat{\rho}$ by more than nine times, from 0.039 to 0.351, and strengthens the significant level of $\hat{\rho}$ from 10% to 5%. The result indicates that a 10% increase in the number of new firms created in the contiguity-weighted neighboring regions is associated with a 3.51% increase in the number of new firms created in the focal region. This result also implies that spatial lag ML may produce downward biased estimation of $\hat{\rho}$ as a consequence of the upward bias on $\hat{\gamma}$. Therefore, a proper handling of the joint endogeneity of the serial- and spatial-lagged dependent variable, as the system GMM has done in this research, is essential for consistent and unbiased estimation of both $\hat{\gamma}$ and $\hat{\rho}$.

Third, the coefficients on four locational variables are now all statistically significant and have the expected signs. This means that a larger local market, a higher level of agglomeration, and more land available for building factories would lead to a higher level of manufacturing start-up activities in the focal region, whereas a higher local wage rate would lead to a lower level of manufacturing start-up activities in the region. These results support the perspective of spatial heterogeneity, in which regional differences in start-up activities can be explained by the presence of substantially different regional

characteristics of the founding locations. In addition, the absolute values of their coefficients increase by a large margin and the order of the magnitudes of these coefficients changes as well. For example, the coefficient on $\ln Agglomeration_{i,t-1}$ increases to 0.598 from 0.067 and 0.066 with the pooling OLS and Spatial lag ML, respectively, resulting in the agglomeration variable having the highest positive elasticity among all explanatory variables. On the other hand, the negative effect of local wage rate becomes much stronger with the system GMM estimation: a 1% increase in local wage rate may lead to a 1.148% decrease in the level of new firm creation activities.

Two key issues for the robustness check exist. The first is the effectiveness of controlling the regional size effect by $\ln Land_{i,t-1}$. To check this effectiveness issue, we construct the size-adjusted dependent variable $\ln(Firm/Land)_{i,t}$ which is the logarithm of one plus (the number of newly created entrepreneurial firms)/(the land size available for building factories) in region i and year t . The system GMM regression based on $\ln(Firm/Land)_{i,t}$, $\ln(Firm/Land)_{i,t-1}$ and $w \cdot \ln(Firm/Land)_{i,t}$ produces qualitatively equivalent results to those reported in the system GMM column of Tables 3. The second robustness issue is whether the results hold with an alternative spatial weighting matrix. To check this, we employ the second-order contiguity-weighting matrix in which the cell (i, j) has value 1 if regions i and j are either next-door or next-door-but-one neighbors, and value zero otherwise for regressions reported in Tables 3. The regression results remain qualitatively unchanged when using this alternative spatial weighting matrix. These robustness testing results are available from the authors upon request.

 Insert Table 3 about here

5 Discussion and conclusion

The attractiveness of a region or a cluster of regions to business start-ups has been regarded as fundamental for long-term regional economic development. Accordingly, there has been a growing research interest in the geography of start-ups and the determinants of the rates of entrepreneurship among both academics and policy makers. Although it is widely acknowledged now that entrepreneurship is a process of collective, network-based activities, the joint temporal-spatial context of entrepreneurship is largely missing in the literature on the determinants of the rates of business start-ups. This research fills in this important gap.

This study takes into consideration all three perspectives existing in the literature which explain regional variations in the level of start-up activities. The first is the spatial heterogeneity in terms of local characteristics of the founding location, which may be conducive or restrictive to entrepreneurial ventures. The second is related to the path dependent processes of entrepreneurial capital accumulation and business start-ups, which indicate that start-up activities in the past can exert a self-augmenting and self-reinforcing impact on start-up activities in the future. The third one refers to spatially dependent nature of business start-ups, meaning that decisive factors for new firm creation such as the accessibility to entrepreneurship capital and competition induced by entrepreneurial activities are not bounded within the home region and entrepreneurship capital can be acquired across regions.

To integrate these three perspectives, the study develops a spatial dynamic panel model which explicitly specifies the temporal and spatial dependences while controlling for the heterogeneity of regional characteristics. Using data on new manufacturing ventures created in sub-national regions of Korea, the empirical results show that both temporal and spatial dependences are significant and operate simultaneously. In addition

to the expected and statistically significant impact of regional characteristics that represent spatial heterogeneity, start-up activities in the past exert a significant effect on the level of current start-up activities, in a way as interpreted in Acs (2006, p. 112): ‘Local policy and previous history (path dependence) determine local entrepreneurial climate, which may be embedded in the local infrastructure, regulation, attitudes, educational policies, networks, technology transfer mechanisms, and so forth.’ The empirical results also forcefully demonstrate that new venture creation in a region can be predicted from new ventures created in its contiguous regions, and an integrated treatment of the joint temporal and spatial endogeneity can significantly improve the estimation results. This reinforces in a substantive way Plummer’s (2010, p. 146) statement that ‘entrepreneurship is indeed a phenomenon of both time and space.’ The comparison across alternative estimators including the pooling regression with OLS, spatial panel regression with maximum likelihood (ML) and the system GMM further confirms that the system GMM seems to be the best available estimator for dealing with the joint endogeneity problem of temporal and spatial dependences.

Two implications of this research for policy makers are worth mentioning. First, considering the geographic connection of entrepreneurial activities across neighboring regions, public policy measures to support the entrepreneurs’ founding activities in less-developed regions need to pay closer attention to fostering network opportunities with prospective entrepreneurs and knowledge sources outside the region, in addition to developing common infrastructure within the related regions. Second, for public policy makers who wish to foster entrepreneurial new firm creation with an aim to boost local economies, they are advised to pay due attention to the joint possibility of spatial and temporal dependences and interactions of entrepreneurial activities that may alter conventional understanding of the relative importance of locational determinants. For

example, while the pooling OLS and Spatial lag ML estimations suggest the size of the local market to be the strongest link between local attractiveness and entrepreneurial activities, the system GMM estimation presents a more comprehensive and context-rich picture, in which all major explanatory variables play their expected roles and the strongest attractor is the level of agglomeration. A correct understanding of this comprehensive and context-rich picture would lead to the formulation and implementation of appropriate public policy measures.

Several limitations of this research should be acknowledged. First, because this research is based on the data of new firms created in a single country, Korea, its findings need to be further verified in the context of other countries for generalization. Second, the measurement for entrepreneurial activities depends on a count variable, i.e., the total number of new firms created in each region-year. This measure may not reflect a ‘net’ level of entrepreneurial activities that takes into account the *exit* rate of entrepreneurial firms in each region-year. Third, because of the data limitation at the fine-grained county- and city-level subnational regions, the research is unable to incorporate, in its empirical settings, all relevant location-specific determinants which spur new firm creation activities. We hope that these limitations will be addressed in near future as more researchers exert efforts in investigating the joint temporal and spatial dependences of entrepreneurial activities.

References

- Acs, Z.J. (2006). New firm formation and the region: Empirical results from the United States. In E. Santarelli (Ed.), *Entrepreneurship, Growth, and Innovation: The Dynamics of Firms and Industries* (International Studies in Entrepreneurship, Vol. 12, pp. 105-133). New York: Springer.
- Acs, Z.J., Armington, C., & Zhang, T. (2007). The determinants of new-firm survival across regional economies: The role of human capital stock and knowledge spillover. *Papers in Regional Science*, 86(3), 367–391. DOI: 10.1111/j.1435-5957.2007.00129.x
- Andersson, M., & Koster, S. (2011). Sources of persistence in regional start-up rates - Evidence from Sweden. *Journal of Economic Geography*, 11(1), 179–201. DOI: 10.1093/jeg/lbp069
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.
- Anselin, L. (1996). The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In M. Fischer, H. Scholten, & D. Unwin (Eds.), *Spatial Analytical Perspectives on GIS* (pp. 111-125). London: Taylor and Francis.
- Anselin, L., Bera, A., Florax, R., & Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, 26(1), 77-104. DOI: 10.1016/0166-0462(95)02111-6
- Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2), 277-297. DOI: 10.2307/2297968
- Arenius, P., & Minniti, M. (2005). Perceptual variables and nascent entrepreneurship. *Small Business Economics*, 24(3): 233-247. DOI: 10.1007/s11187-005-1984-x
- Armington, C., & Acs, Z.J. (2002). The determinants of regional variation in new firm formation. *Regional Studies*, 36(1), 33-45. DOI: 10.1080/00343400120099843

- Arthur, W.B. (1994). *Increasing Returns and Path Dependence in the Economy*. University of Ann Arbor: Michigan Press.
- Audretsch, D.B., & Feldman, M.P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, 86(3), 630-640. Stable URL: <http://www.jstor.org/stable/2118216>
- Audretsch, D.B., & Keilbach, M.C. (2004). Entrepreneurship capital and economic Performance. *Regional Studies*, 38(8), 949-959. DOI: 10.1080/0034340042000280956
- Audretsch, D.B., Keilbach, M.C., & Lehmann, E. (2006). *Entrepreneurship and Economic Growth*. Oxford: Oxford University Press.
- Baltagi, B.H. (2008). *Econometric Analysis of Panel Data*. 4th edition, Chichester: John Wiley and Sons.
- Baum, J.A.C., & Sorenson, O. (2003). (Eds) *Geography and Strategy*. *Advances in Strategic Management*, Volume 20, Amsterdam: Elsevier.
- Blundell, R., & Bond, S.R. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143. DOI: 10.1016/S0304-4076(98)00009-8
- Boschma, R.A., & Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, 6(3), 273-302. DOI: 10.1093/jeg/lbi022.
- Bosma, N., Aces, Z., Autio, E., Coduras, A., & Levie, J. (2009). *Global Entrepreneurship Monitor 2008 Executive Report*. Babson Park, MA: Babson College.
- Breitenecker, R. J., & Hams, R. (2010). Dealing with spatial heterogeneity in entrepreneurship research. *Organizational Research Methods*, 13(1), 176-191. DOI: 10.1177/1094428109338871

- Brixy, U., & Grotz, R. (2007). Regional patterns and determinants of birth and survival of new firms in Western Germany. *Entrepreneurship and Regional Development*, 19(4), 293-312. DOI: 10.1080/08985620701275510
- Brueckner, J.K. (2011). Lectures on Urban Economics. Cambridge: The MIT press.
- Brush, C.G., & Vanderwerf, P.A. (1992). A comparison of methods and sources for obtaining estimates of new venture performance. *Journal of Business Venturing*, 7(2), 157-170. DOI: 10.1016/0883-9026(92)90010-O
- Cairncross, F. (1997). The Death of Distance. Boston: Harvard Business School Press.
- David, P.A. (1985). Clio and the economics of QWERTY. *The American Economic Review*, 75(2), 332-337. Stable URL: <http://www.jstor.org/stable/1805621>
- Elhorst, J.P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26(3), 244-268. DOI: 10.1177/0160017603253791
- Elhorst, J.P. (2005). Unconditional maximum likelihood estimation of dynamic models for spatial models. *Geographical Analysis*, 37(1), 85-106. DOI: 10.1111/j.1538-4632.2005.00577.x
- Folta, T.B., Cooper, A.C., & Bail, Y. (2006). Geographic cluster size and firm performance. *Journal of Business Venturing*, 21(2), 217-242. DOI: 10.1016/j.jbusvent.2005.04.005
- Fornahl, D. (2003). Entrepreneurial activities in a regional context. In D. Fornahl, & T. Brenner (Eds.), *Cooperation, Networks, and Institutions in Regional Innovation Systems* (pp. 38-57). Cheltenham: Edward Elgar.
- Fredin, S. (forthcoming). New perspectives on innovative entrepreneurship and path dependence – A regional approach. *International Journal of Entrepreneurial Venturing*, in press.

- Fritsch, M., & Falck, O. (2007). New business formation by industry over space and time: A multidimensional analysis. *Regional Studies*, 41(2), 157-172. DOI: 10.1080/00343400600928301
- Fritsch, M., & Mueller, P. (2007). The persistence of regional new business formation-activity over time - assessing the potential of policy promotion programs. *Journal of Evolutionary Economics*, 17(3), 299-315. DOI: 10.1007/s00191-007-0056-6
- Fritsch, M., & Mueller, P. (2008). The effect of new business formation on regional development over time: The case of Germany. *Small Business Economics*, 30(1), 15-29. DOI: 10.1007/s11187-007-9067-9
- Fujita, M., & Thisse, J.-F. (2002). Economics of agglomeration: Cities, industrial location, and regional growth. London and New York: Cambridge University Press.
- Graf, H. (2011). Gatekeepers in regional networks of innovators. *Cambridge Journal of Economics*, 35(1), 173-198.
- Head, K., Ries, J., & Swenson, D. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investment in the United States. *Journal of International Economics*, 38(3/4), 223-247. DOI: 10.1016/0022-1996(94)01351-R
- Holcombe, R.G. (2007). Entrepreneurship and economic growth. In B. Powell (Ed.), *Making Poor Nations Rich: Entrepreneurship and the Process of Economic Development*. Stanford: Stanford University Press.
- Hsiao, C. (2003). *Analysis of Panel Data*. 2nd edition. Cambridge: Cambridge University Press.
- Johnson, P., & Parker, S. (1996). Spatial variations in the determinants and effects of firm births and deaths. *Regional Studies*, 30(7), 679-688. DOI: 10.1080/00343409612331349968

- Keeble, D., & Walker, S. (1994). New firms, small firms and dead firms: Spatial patterns and determinants in the United Kingdom. *Regional Studies*, 28(4), 411-427. DOI: 10.1080/00343409412331348366
- Kukenova, M., & Monterio, J.-A. (2008). Spatial dynamic panel model and system GMM: A Monte Carlo investigation. MPRA Paper No. 14319, University Library of Munich.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483-499. Stable URL: <http://www.jstor.org/stable/2937739>
- Lasch, F., Robert, F., & Roy, F.L. (2013). Regional determinants of ICT new firm formation. *Small Business Economics*, 40(3), 671-686. DOI: 10.1007/s11187-011-9382-z
- Maitland, E., Rose, E.L., & Nicholas, S. (2005). How firms grow: Clustering as a dynamic model of internationalization. *Journal of International Business Studies*, 36(4), 435-451. DOI: 10.1057/palgrave.jibs.8400140
- Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, 6(4), 395-437. DOI: 10.1093/jeg/lbl012
- Minniti, M. (2005). Entrepreneurship and network externalities. *Journal of Economic Behavior and Organization*, 57(1), 1-27. DOI: 10.1016/j.jebo.2004.10.002
- Nijkamp, P. (2003). Entrepreneurship in a modern network economy. *Regional Studies*, 37(4), 395-405. DOI: 10.1080/0034340032000074424
- O'Donnell, A., Gilmore, A., Cummins, D., & Carson, D. (2001). The network construct in entrepreneurship research: a review and critique. *Management Decision*, 39(9), 749 - 760. DOI: 10.1108/EUM0000000006220

- Oort, F.G. van, & 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(2), 263-290. DOI: 10.1007/s00168-004-0195-8.
- Pijnenburg, K., & Kholodilin, K.A. (2014). Do regions with entrepreneurial neighbours perform better? A spatial econometric approach for German regions. *Regional Studies*, 48(5), 866-882. DOI: 10.1080/00343404.2012.697143
- Plummer, L.A. (2010). Spatial dependence in entrepreneurship research: Challenges and methods. *Organizational Research Methods*, 13(1), 146-175. DOI: 10.1177/1094428109334199
- Porter, M.E. (1998). Clusters and the new economics of competition. *Harvard Business Review*, 76(6), 77-90.
- Reynolds, P.D., Bygrave, W.D., Autio, E., & Hay, M. (2002). Global entrepreneurship monitor 2002 summary report. Babson Park, MA: Babson College.
- Reynolds, P.D., Miller, B., & Maki, W.R. (1995). Explaining Regional Variation in Business Births and Deaths: U.S. 1976-88. *Small Business Economics*, 7(5), 389-407. DOI: 10.1007/BF01302739
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158. DOI: 10.1111/j.1468-0084.2008.00542.x
- Schienstock, G. (2007). From path dependency to path creation: Finland on its way to the knowledge-based economy. *Current Sociology*, 55(1), 92-109. DOI: 10.1177/0011392107070136
- Schwab, K. (2009). The Competitiveness Report 2009-2010. Geneva, Switzerland: World Economic Forum.
- Shane, S. (2003). A General Theory of Entrepreneurship: The Individual-opportunity Nexus. Cheltenham: Edward Elgar.

- Smith, H.L., Glasson, J., & Chadwick, A. (2005). The geography of talent: entrepreneurship and local economic development in Oxfordshire. *Entrepreneurship and Regional Development*, 17(6), 449-478. DOI: 10.1080/08985620500247819
- Stam, E., & Lambooy, J. (2012). Entrepreneurship, knowledge, space, and place: Evolutionary economic geography meets Austrian economics. In D.E. Andersson (Ed.), *The Spatial Market Process* (Advances in Austrian Economics, Volume 16, pp. 81-103). Bingley: Emerald.
- Steyaert, C., & Katz, J. (2004). Reclaiming the space of entrepreneurship in society: Geographical, discursive and social dimensions. *Entrepreneurship and Regional Development*, 16(3), 179-196. DOI: 10.1080/0898562042000197135
- St-Jean, E., & Audet, J. (2012). The role of mentoring in the learning development of the novice entrepreneur. *International Entrepreneurship and Management Journal*, 8(1), 119-140. DOI: 10.1007/s11365-009-0130-7
- Stuart, T., & Sorenson, O. (2003). The geography of opportunity: Spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy*, 32(2), 229-253. DOI: 10.1016/S0048-7333(02)00098-7
- Sun, L., Hong, E., & Li, T. (2010). Incorporating technology diffusion, factor mobility and structural change into cross-region growth regression: An application to China. *Journal of Regional Science*, 50(3), 734-755. DOI: 10.1111/j.1467-9787.2009.00623.x
- Trettin, L., & Welter, F. (2011). Challenges for spatially oriented entrepreneurship research. *Entrepreneurship and Regional Development*, 23(7-8), 575-602. DOI: 10.1080/08985621003792988

- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), 25-51. DOI: 10.1016/j.jeconom.2004.02.005
- Yu, J., de Jong, R., & Lee, L. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146 (1), 118-134. DOI: 10.1016/j.jeconom.2008.08.002
- Zucker, L. G., Darby, M. R., & Brewer, M. B. (1998). Intellectual human capital and the birth of U.S. biotechnology enterprises. *The American Economic Review*, 88(1), 290-306. Stable URL: <http://www.jstor.org/stable/116831>

Table 1 Moran's I Test for spatial autocorrelation

$\ln Firm_{it}$	Moran's I	Z-statistics
$t = 2000$	0.350***	8.418
$t = 2001$	0.315***	7.603
$t = 2002$	0.287***	6.936
$t = 2003$	0.276***	6.662
$t = 2004$	0.333***	8.028

Notes. *** denotes significance level at 1%

Table 2 Descriptive statistics and correlation matrix ^{a,b}

	Label	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7
1	$\ln Firm_{it}$	2.69	1.56	0.00	6.38	1.00						
2	$\ln Firm_{i,t-1}$	2.54	1.52	0.00	6.01	0.90***	1.00					
3	$\ln GRP_{i,t-1}$	13.29	1.91	5.74	17.54	0.76***	0.76***	1.00				
4	$\ln Agglomeration_{i,t-1}$	9.09	0.94	6.81	11.17	0.53***	0.54***	0.60***	1.00			
5	$\ln Wage_{i,t-1}$	0.26	0.32	-0.79	1.27	0.45***	0.44***	0.65***	0.24***	1.00		
6	$\ln Land_{i,t-1}$	0.24	1.22	-2.93	3.39	0.08**	0.08**	0.08***	-0.39***	0.08***	1.00	
7	$W_1 \cdot \ln Firm_{it}$	2.82	1.01	0.35	6.08	0.50***	0.49***	0.48***	0.38***	0.37***	-0.10***	1.00

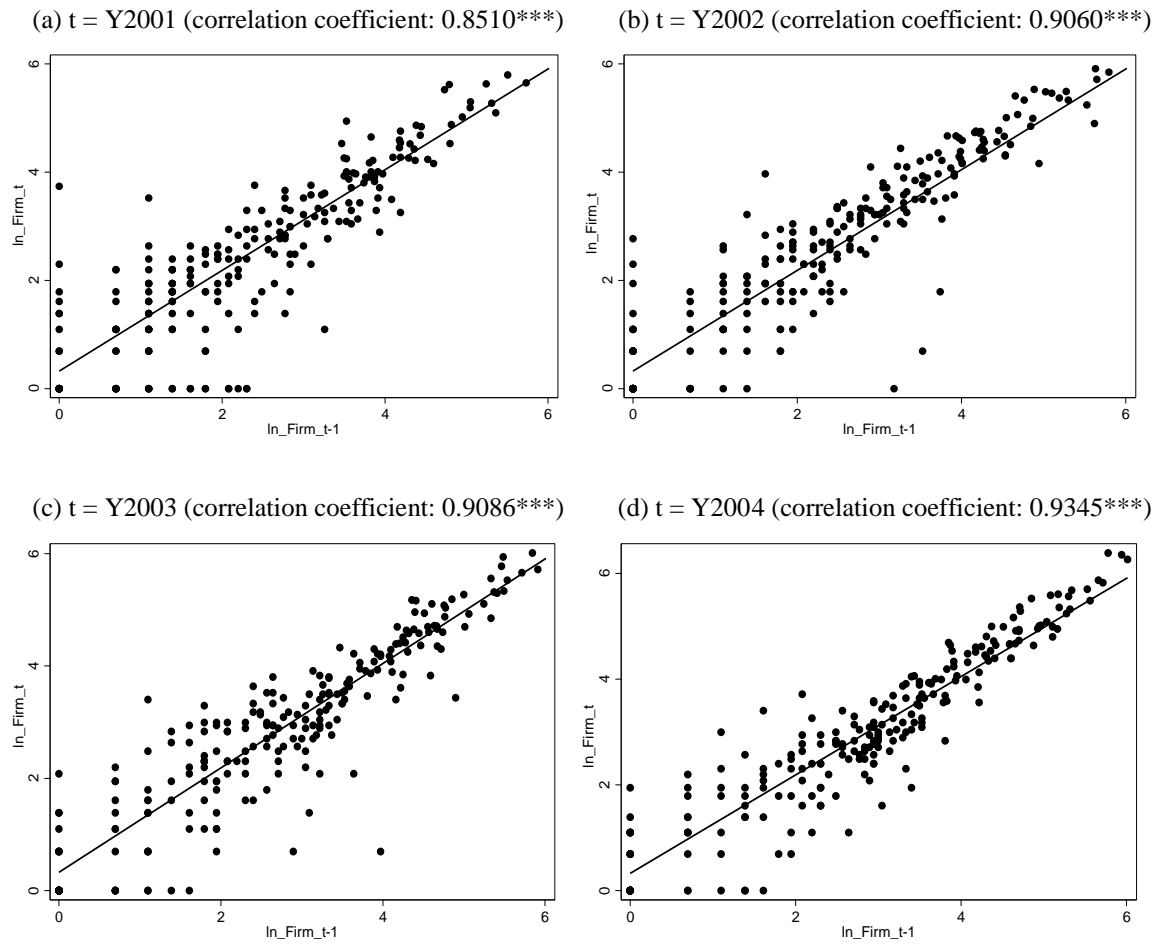
Notes. a. N = 936 (Due to $\ln Firm_{i,t-1}$, the number of observation reduces from 1,170 to 936). b. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3 Estimation results

Variables	Pooling OLS	Spatial lag ML	System GMM
$\ln Firm_{i,t-1}$	0.763 [0.023]***	0.761 [0.023]***	0.414 [0.207]**
$\ln GRP_{i,t-1}$	0.126 [0.025]***	0.122 [0.024]***	0.233 [0.127]*
$\ln Agglomeration_{i,t-1}$	0.067 [0.036]*	0.066 [0.035]*	0.598 [0.341]*
$\ln Wage_{i,t-1}$	0.023 [0.098]	0.013 [0.098]	-1.148 [0.582]**
$\ln Land_{i,t-1}$	0.010 [0.022]	0.015 [0.022]	0.389 [0.192]**
$\mathbf{W} \cdot \ln Firm_{it}$		0.039 [0.024]*	0.351 [0.151]**
Constant	-1.570 [0.293]***	-1.598 [0.292]***	-7.862 [2.618]***
Adj-R ² / Log likelihood	0.829	-912.086	
Moran's I (residuals, <i>p</i> -value)	(0.351)		
LM _{error} (<i>p</i> -value)	(0.445)		
Robust LM _{error} (<i>p</i> -value)	(0.295)		
LM _{lag} (<i>p</i> -value)	(0.000)		
Robust LM _{lag} (<i>p</i> -value)	(0.000)		
Wald test (<i>p</i> -value)		(0.096)	
LR test (<i>p</i> -value)		(0.096)	
LM test (<i>p</i> -value)		(0.092)	
No. of observations	936	936	936
Number of Instruments			35
Hansen <i>J</i> test (<i>p</i> -value)			(0.881)
Difference-Hansen tests (<i>p</i> -value)			
All system GMM instrument			(0.768)
Those based on lagged number of firm creation only			(0.189)
AR(1) test in differences (<i>p</i> -value)			(0.003)
AR(2) test in differences (<i>p</i> -value)			(0.594)

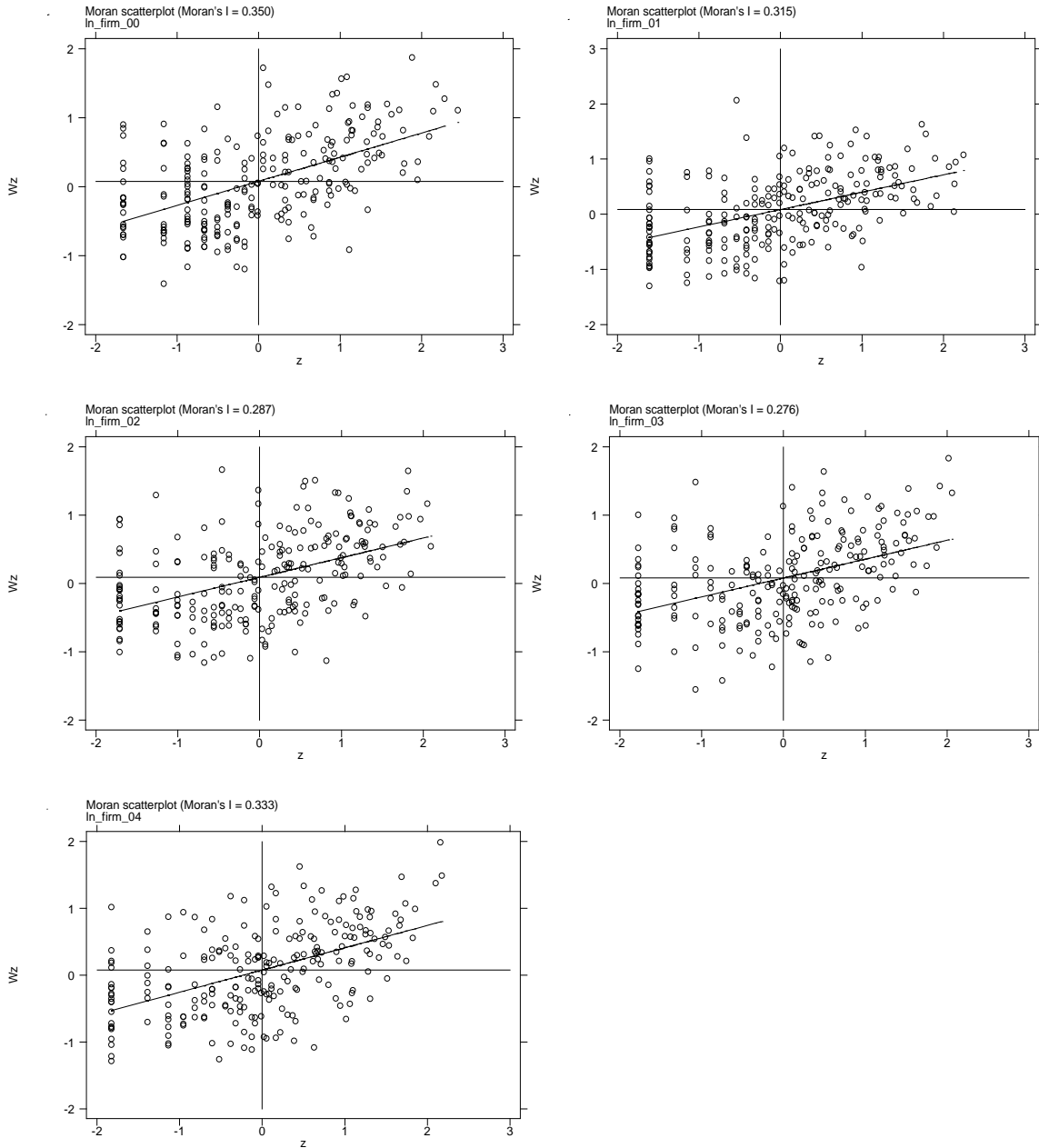
Notes. Numbers in [] and () are standardized errors and *p*-values respectively. *, ** and *** denotes significance levels at 10%, 5%, and 1%, respectively. **W** is a first-order contiguity-weighted matrix. Industrial complex dummies and yearly dummies are not reported.

Fig. 1 Temporal dynamics of Korean entrepreneurship: Scatter plots



Notes. *** denotes significance level at 1%

Fig. 2 Spatial dependence of Korean entrepreneurship: Moran scatter plots



Note. All Moran's-I coefficients are significance at 1% level (Table 1).