Pettersson, Lars; Sjölander, Pär; Widell, Lars M

Conference Paper

Firm formation in rural and urban regions explained by demographical structure

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Pettersson, Lars; Sjölander, Pär; Widell, Lars M (2010) : Firm formation in rural and urban regions explained by demographical structure, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/119018

Firm Formation in Rural and Urban Regions Explained by Demographical Structure

Lars Pettersson
The Swedish Board of Agriculture and Jönköping International Business School, 551 82 Jönköping, Sweden, e-mail: lars.pettersson@sjv.se, Phone: +4636155119

Pär Sjölander
The Swedish Board of Agriculture and Jönköping International Business School, 551 82 Jönköping, Sweden, e-mail: Par.Sjolander@sjv.se, Phone: +4636156144

Lars M Widell (corresponding author)
The Swedish Board of Agriculture, Jönköping International Business School and Örebro University, 551 82 Jönköping, Sweden, e-mail: larsm.widell@sjv.se, Phone: +4636155026

Abstract

In this paper we focus on the question of how population demography influences firm formation. In particular, we focus on this question with respect to the presence of differences between urban and rural areas, and if the demographical structure is a significant factor that can explain regional variation in firm formation? In the empirical analysis we perform a cross-sectional analysis where we use data on municipalities in Sweden. Furthermore, we make a distinction between the two types of regions in the Swedish economy (rural and urban) in order to analyze how these areas differs from each other with respect to our specific question of how population demography influences firm formation in rural and urban areas. By the use of spatial regression models the results show that households in the age where we can assume they have small children have a negative effect on the propensity of firm formation, households that are in the retirement age have a positive effect on the propensity of firm formation. We also find that rurality have a positive effect on firm formation. Our results correspond to other studies in the area with respect to a positive correlation between firm formation and age.

Paper to be presented at the 50th Anniversary European Congress of the Regional Science Association International: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19 - 23 August 2010 - Jonkoping, Sweden

1 Introduction

New firm formation is usually considered as one important aspect of entrepreneurship that also is vital for the dynamics of the economy and long term stimulation for economic growth. The role of entrepreneurs and entrepreneurship, as explanation to economic progress, was highlighted 1912 by Joseph Schumpeter, and later by Baumol (1968). Research on entrepreneurship has received a lot of attention during the last decades, and various policy measures are used in order to stimulate new firm formation. One reason for this interest in new firm formation, and also the role of small and medium sized firms in the economy, is the way these firms are important for job creation.

There is a substantial amount of studies on regional variation of firm formation. In particular, empirical analysis can be found that use data from the 1980s and onwards, and studies have been made for many different countries. An overview of this literature (see for example the special issues of Regional Studies from 1994 and from 2004) shows that there are significant regional variations in new firm formation, and that a number regional factors contributes to explain differences in new firm formation. Such factors are unemployment, human capital/educational level, differences in income, etc.

Entrepreneurship and new firm formation is not only tied to economical factors, also social factors –that varies across regions– can be assumed to contribute to explain spatial differences. In the literature, it is common to make distinction between different types of entrepreneurship. Westlund (200X) distinguish between four types of entrepreneurship: (1) economic/technical, (2) social/civil, (3) academic and (4) political/administrativ. Social factors have often been analyzed in network context, and also with respect to social norms and institutional aspects. The importance of social factors has, for example, been in focus of attention in analysis of how the so-called knowledge based economy (industries with labor that has high education) has emerged (Westlund, 2004).

The theoretical advancement in the field of the new economic geography and endogenous growth theory has contributed to a deeper understanding of spatial variation in economic development. Studies in this field frequently show how urbanization is an important factor that stimulates economic growth. Increasing returns to scale in production, in particular tied to external conditions related to the size of markets, explains how urbanization can drive economic growth.

Accordingly, it is reasonable to presume that new firm formation should be comparatively strong in urban regions due to economies of agglomeration. At the same time, previous studies (for example Armington and Acs, 2002) show that the regional average size of firms is likely to have a negative influence on new firm formation. This means that regions that are characterized by small scale economies can be assumed to perform relatively strong with respect to new firm formation. This means that it is not self evident whether or not rural areas should be expected to have a lower level of new firm formation compared to urban areas. This question is also what we have in focus in this study, and the empirical analysis employs Swedish data.

One important social characteristic of the rural areas in Sweden is the long term changes in the demographic structure. A stylized and very general view of the rural areas in Sweden show a

weak population growth compared to urban areas (Glesbygdsverket, 2008). In particular young cohorts (18-24 years of age) tend to migrate from rural areas toward cities with universities and university-colleges. This structural "imbalance" between rural and urban areas is often notified and seen as a potential "drain-problem" for rural areas. In relation to this, we address the importance of demographical structure with respect to new firm formation.

In this paper we focus on the question of how a social factor such as population demography influences firm formation. In particular, we focus on this question with respect to the presence of differences between urban and rural areas, and if the demographical structure is a significant factor that can explain regional variation in firm formation? In the empirical analysis we perform a cross-sectional analysis where we use data on municipalities in Sweden. Furthermore, we make a distinction between the two types of regions in the Swedish economy (rural and urban) in order to analyze how these areas differs from each other with respect to our specific question of how population demography influences firm formation in rural and urban areas.

The purpose of this study is to analyze how a social factor such as population demography influences firm formation. Furthermore, is there a difference between urban and rural areas? The empirical analysis is conducted by the use of data on local level (municipalities) in Sweden. We employ a cross-sectional dataset for approximately 285 Swedish municipalities in 2004.

2 Background

Geographical differences in industrial specialization is a well documented and analyzed area of research. The theory of the new economic geography has grown vastly, in particular from the 1970s and onwards. Seminal contributions, like Porter (1990), Krugman (1991) and Fujita, Krugman and Venables (1999), has had a lot of impact on how this area of research have been developed. The research in this field states very clearly that we should expect differences in regional performance and development according to a number of factors that characterize any region. Economic activities and factors of production are not evenly distributed in space, and distances (transportation and commuting costs) between markets are also important to take into consideration in analysis of regional development. The size of markets – in terms of for example purchasing power – is also a factor that determines the potential for a number of industries to benefit from increasing economies of scale in production (which is present in firms that have fixed costs).

In addition to production prerequisites that are determined by regional characteristics, we also find households (the demand side) preferences for variety. The importance of considering how product variety can explain why a substantial amount of all markets can be assumed to be characterized by monopolistic competition (of different degrees) was recognized by Lancaster (1966). One key findings, which the notation of households preferences for variety gives, is that a market that is able to supply a great variety in products on a local or regional market will also be likely to attract households. This means that a self reinforcing endogenous growth mechanism can be established in urban regions, driven by the size of the market, that attracts both firms (that can benefit from external economies tied to the market size) and households (that are attracted by a large variety in the supply of goods and services). Accordingly, this field of research stipulates that urbanization is one important factor that can explain economic growth (Quigley, 1998).

Without doubt, urbanization can be assumed to be one of the most important factors explaining growth and development of regions. There are, however, also other factors that drives these processes. Some example of those factors are access to factors of production, and industrial specialization that evolves from vertical and horizontal integration. There are also other non-economical factors that are relevant to include in the analysis of regional growth and performance. One such perspective is social factors and characteristics of the civil society that also differs between regions. In this study we are particularly focused on the demographical structure with respect to age cohorts.

There is an extensive literature written that connects the civil society and so-called social capital with regional development and business activities. Most of this literature is case oriented tied to different places etc. There are, however, theoretical and comprehensive studies to be found in this field of research by Putnam (1993 and 2000) and Florida (2002, 2003, 2005 and 2008). The studies by Putnam and Florida are very different in many respects, both in terms of theoretical and empirical approaches. Nevertheless, both authors put a significant focus on how social factors influence regional economic development.

In a brief and very simplified summary, we can conclude that Putnam puts a great emphasis on the way a community with strong social networks and homogenous norms and values stimulate economic performance in a region. According to Putnam, *trust* between actors in a local or regional society is a key element for collaboration and entrepreneurship. Trust also reduces transaction costs and is a fundament for a "good" community that also attract households. The question of "good" institutions is also something that usually connects with the "good" community in this type of framework. Empirical analysis of social capital, based on Putnam´s theoretical standpoint often use various types of proxies that are assumed to reveal trust such as membership in different non-profit associations, participation in elections, etc.

Florida puts a focus on what he refers to as the *creative class*. In his book from 2002, Richard Florida made a classification of the economy (the US economy) and made a calculation that showed that around 30 percent of the population in the US economy could be regarded as belonging to the creative class. Florida's argument is then that analysis of regional development that also includes the consideration of the knowledge society, should focus on question of where the *creative class people* chose to live. The creative class is not evenly distributed in space, and one interesting question to analyze is therefore the factors that influence this population groups such as talent, the choice of living. Florida did also construct an index composed by characteristics like technology and tolerance, which he assumes are reflections of factors of attraction for the creative class. According to Florida, we should expect the creative class to be attracted by dense populated markets that also facilitates job mobility, a regional diversity in culture, lifestyles, ethnicity, large number of meeting places (that also are diversified), places with strong and unique identity (and/or ability to form its own identity) and a large variety in services and supply of leisure activities etc. Florida also suggest that there is a causality link from since he stipulate that creative capital is attracted by places that scores high on the index he have formulated (Florida, 2002).

As seen above there are several differences between Putnam and Florida´s approaches. One such difference is that while Putnam views "social capital" as a fundamental driving mechanism for regional development Florida focus on "creative capital". In consequence, these two views results in quite different empirical approaches with respect to studies of how social factors

influence economic performance. There are also limitations with both of these two views of how social factors can explain economic development. The explanatory variables that can be used, based on Putnam´s theory, are very much reflections of "social capital", which not necessary is the same as a stock of social capital. There is no general widely accepted definition of such stock of social capital. Florida´s index, which could be assumed to reflect a creative capital stock (but it's not really a stock) becomes very much correlated with urbanization. In addition, in an economy like the Swedish, regions with a high share of refugee immigration, or one single plant of for example Ericsson or ABB scores very high. However, such regions are not necessary very diversified and entrepreneurial activities such as new firm formation and/or innovation activities – which often are seen as a reflection of creativity – are not always correlated with a high score on Florida´s index. This does not mean that Florida´s index does not work out well for the US economy and other countries.

Many studies that focus on how social factors have influence on economic performance are often using the notation of civil society as a starting point for analysis. Civil society is often defined as the propensity of people to voluntary associate with organizations in the civil society that are based on no personal profit (or similar definition). This is sometimes referred to as "the third sector" that also can be regarded as a form of capital. This view appears to be very interesting, however, there is still the problem that we cannot find a clear definition of the stock of the third sector.

In Sweden, the government have formulated a strategy in purpose to ensure a sustainable development. This strategy (Regeringens skrivelse 2005/06:126) includes environmental, economical and social perspectives. In the formulation of the strategy the government has had the ambition to be as precise as possible with respect to defining the objectives for the policy on sustainable development. This means that a number of indicators have been presented that are assumed to reflect social cohesion. These indicators are: risk of poverty, demographic dependency ratio, distribution of income, children at risk of poverty, financial crisis, people born abroad in financial crisis, regional demographic change, long term unemployment, people born abroad (employment), sickness absenteeism, fertility, parental leave, equal opportunity wages, equal opportunity managers, electoral participation, people born abroad electoral participation, trust in the media, school security, basic requirements of the educational system, housing overcrowding, loneliness, consumption of culture and computers and broad band. In total there are 25 indicators in the Swedish strategy for sustainable development with respect to social cohesion. The indicators are measured very differently and data are collected from registers and by means of surveys. It is not possible to merge all the indicators into one in order to construct a stock measure of social cohesion and it would also be a delicate task to make a choice of how to weight these 25 perspectives of social factors against each other.

Based on the overview presented above our choice of social factors for explaining new firm formation across Swedish regions is to focus on the variation in the structure of population age cohorts. Since we cannot find one definition of social capital that is used with a broad consensus among different scholars, we have chosen to focus on population structure based on the assumption that this structure is one interesting aspect of social factors that also differs between regions. Our interest in the analysis of how rural areas differs from urban areas also serves as an extra motivation for our choice since we know from other studies (Glesbygdsverket, 2007) that young people tend to migrate from rural areas to urban areas. In relation to this observation it is of interest to understand if such change in population structure has a negative impact on new firm formation.

Our overview also shows that a number of social perspectives can be assumed to correlate with urbanity and/or population agglomeration. We will utilize this observation by make comparison between urban and rural areas. We also conclude that it is not possible to elaborate with the 25 Swedish dimensions of social perspectives of sustainable development. We do not have access to data on all these variables/indicators, and it appears to be difficult to decide on how these perspectives should be weighted against each others. We also rest our empirical analysis on the assumption that social development is strongly tied to people, and that the age structure is one important element of this.

3 The Swedish Economy and the Urban-Rural Perspective

In this study we use a definition of rural and urban areas that has been applied in numerous other studies of the Swedish economy (SOU 2006:1, Johansson & Klaesson 2008, Jordbruksverket 2009 RA09:3), where municipalities in Sweden are classified into four different groups, municipality type (MT) 1, 2, 3, and 4. (MT 1) metropolitan areas, (MT 2) urban areas, (MT 3) rural areas/countryside, and (MT 4) sparse populated rural areas. This definition of regions builds on analysis that has been carried out at the Swedish Board of Agriculture. The four types of areas are defined as follows:

- Metropolitan areas (MT 1): Includes municipalities where 100 percent of the population lives within cities or within a 30 km distance from the cities. Using this definition, there are three metropolitan areas in Sweden: Stockholm, Gothenburg and Malmö.
- Urban areas (MT 2): Municipalities with a population of at least 30 000 inhabitants and where the largest city has a population of 25 000 people or more. Smaller municipalities that are neighbours to these urban municipalities will be included in a local urban area if more than 50 percent of the labour force in the smaller municipality commutes to a neighbour municipality. In this way, a functional-region perspective is adopted.
- Rural areas/countryside (MT 3): Municipalities that are not included in the metropolitan areas and urban areas are classified as rural areas/countryside, given they have a population density of at least 5 people per square kilometre.
- Sparse populated rural areas (MT 4): Municipalities that are not included in the three categories above and have less than 5 people per square kilometre.

The above categorization of the Swedish municipalities into the four groups of regions is graphically displayed in the map below.

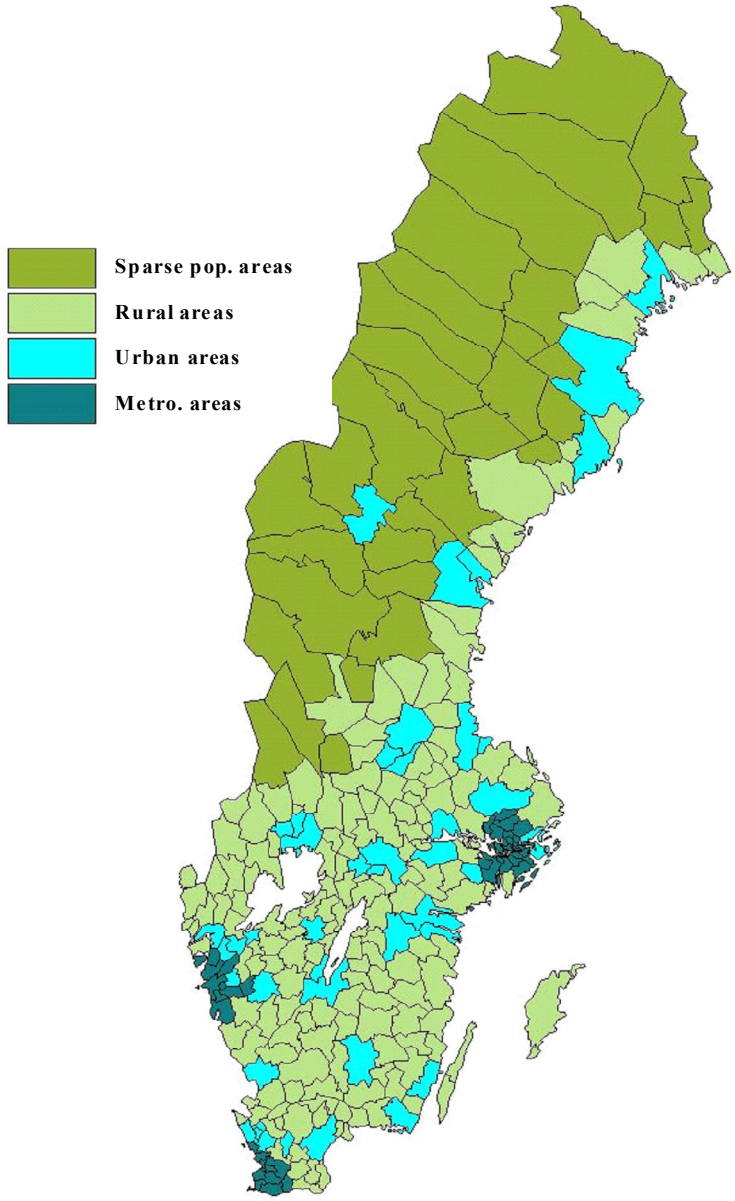

Figure 3.1: *Swedish municipalities classified into (MT 1) metropolitan areas, (MT 2) urban areas, (MT 3) rural areas/countryside, and (MT 4) sparse populated rural areas*

Furthermore, in this study we add the two categories of rural areas (rural areas/countryside and sparse populated areas) into one representing the rural part of Sweden, and the two categories urban areas and metropolitan areas into one group representing the urban part of the economy.

There are a number of substantial differences in the development over time for the rural and urban areas in Sweden. When we compare the population change in the two types of areas between the years 2000 and 2007 we find that urban areas have increased their population by 6 percent at the same time as rural areas have declined their population by 1 percent. During these years, 2000-2007, the number of firms in Sweden increased by 16 percent. It is interesting to notice that the number of firms increased by 20 percent in the urban areas while the corresponding growth rate in rural areas was 10 percent. On the one hand we can conclude that the overall total growth of firms was significantly stronger in urban areas compared to rural areas. But on the other hand, if we compare the relative growth of the number of firms

with respect to population change, the rural areas exhibit a relatively stronger development compared to the urban areas (table 2.1 below). Consequently, the entrepreneurship is relatively strong in the rural areas.

Table 3.1: *Population, population change and firm growth in rural and Urban areas in Sweden 2000-2007*

	Population 2000	Population 2007	Population change 2000-2007	Growth in numbers of firms 2000-2007
Urban	5 653 570	5 976 389	6%	20%
Rural	3 229 222	3 206 538	-1%	10%
Total	8 882 792	9 182 927	3%	16%

The strongest growth performance in the rural areas during the years 2000-2007 is found in the sector of agriculture, forestry and hunting. One can notice that the number of firms increased by approximately 8 percent in the rural areas during these years. Other sectors that exhibited a relative growth during the same time frame in the rural areas were real estate, private services, health care and construction. The relatively strongest growth of firms in urban areas during the years 2000-2007 was in the real estate sector, private services, health care and agriculture/forestry/hunting. This means that the sector agriculture/forestry/hunting exhibited strong growth with respect to firm formation both in urban and rural areas, which is illustrated in the figure below.

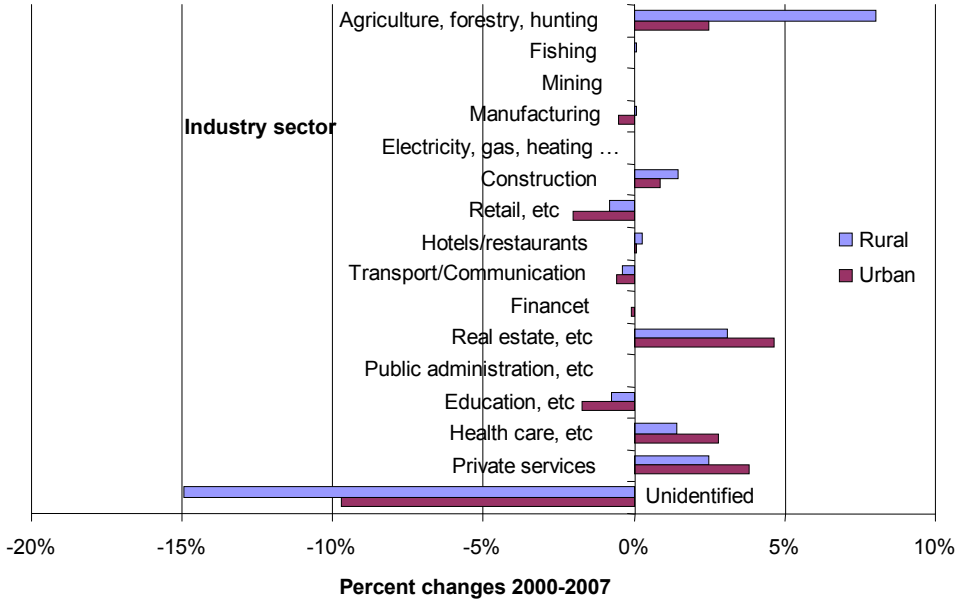

Figure 3.2: *Industry growth 2000-2007 in different industrial sectors in rural and urban areas in Sweden*

Based on an earlier study (Swedish Board of Agriculture, 2005) the firms (or more specifically, the work places) in the agriculture industry as well as forestry, hunting and fishing are characterized by a high share of firms that do not have any employees. More than 90 percent of

the firms in these industries do not have any employees. Table 3.2 below displays the structure of work places in rural areas in Sweden in the year 2002. Around 30 percent of all firms (work places) in the country were located in the rural areas in 2002, and around one third of the rural firms were oriented towards agriculture. This means that almost 10 percent of all firms in Sweden are agricultural firms.

Table 3.2: *Work places in rural areas in Sweden 2002*

Industry sector	Number of work places	Work places without employees	Industry sector	Number of work places	Work places without employees
Food and beverage	1 180	59%	Finance and business services	25259	80%
Agriculture	87 503	93%	Education and research	3189	38%
Forestry, hunting and fishing	14 506	90%	Health care	6090	33%
Manufacturing and mining	15 699	61%	Private services	14247	66%
Energy, water and waste	1 438	63%	Public administration	470	19%
Construction	17 481	66%	Not specified	36322	95%
Retail and communication	31 993	63%	*Total*	*255 377*	*80%*

Source: Statistics Sweden's Business Register (2002).

Based on earlier studies and from the stylized facts we have presented in the first part of this study we are able to make some notations that are of interest for our study:

- Small firms are important with respect to the contribution on job creation. This has been shown both in international and in Swedish studies (an observation valid both in booms and in recessions).
- Other studies show that small firms host larger numbers of jobs (around 67 percent in 1990-93) compared to their GDP-contribution (around 50 percent in 1990-93).
- The growth of firms (in relation to population change) is relatively stronger per capita in rural areas compared to urban areas.
- The agricultural sector host a substantial fraction of all firms and the sector is characterized as a growth sector both in rural and urban areas.

4 Empirical analysis

The model used in this empirical part can be written as:

$$\left(\frac{\text{start}-\text{ups}}{\text{population}}\right)_{i,t} - \left(\frac{\text{start}-\text{ups}}{\text{population}}\right)_{i,t-1} = \left(\frac{\text{small enterpr.}}{\text{all enterpr.}}\right)_{i,t} + \left(\frac{\text{employment}}{\text{population}}\right)_{i,t} + \left(\frac{\text{population cohort}}{\text{total populaton}}\right)_{i,t}$$

We also "throw in" a dummy variable indicating whether the region is a rural one or an urban one (URB = 1 if urban region). This can be written somewhat more condensed,

$$\left(Y_{i,t} - Y_{i,t-1}\right) = \text{small}_{i,t} + \text{emprate}_{i,t} + \text{cohortrate}_{i,t} + URB + \text{error}_{i,t} \qquad \text{(eq. 1)}$$

where $(Y_{i,t} - Y_{i,t-1})$ is called ***ERATED*** in the regression results below. This model is regressed using three different specifications: i) OLS, ii) spatial lag and iii) spatial error. The reason for this is that there are problems with spatial autocorrelation. The OLS is calculated according to equation 1, using robust standard errors. The spatial lag model is calculated according to,

$$\left(Y_{i,t} - Y_{i,t-1}\right) = X\left(Y_{i,t} - Y_{i,t-1}\right) + \text{small}_{i,t} + \text{emprate}_{i,t} + \text{cohortrate}_{i,t} + URB_i + \text{error}_{i,t} \text{ (eq. 2)}$$

where X is a spatial weight matrix, and the spatial error model is calculated as

$$\begin{aligned} \left(Y_{i,t} - Y_{i,t-1}\right) &= \text{small}_{i,t} + \text{emprate}_{i,t} + \text{cohortrate}_{i,t} + URB_i + \text{error}_{i,t} \\ \text{error}_{i,t} &= X\left(\text{error}_{i,t}\right) + \eta_{i,t} \end{aligned} \qquad \text{(eq. 3)}$$

Table ?.? below displays the results from the tests for spatial autocorrelation.

Table ?.? Testing for spatial autocorrelation

Cohort	Spatial error (Moran's I)	Spatial lag (Robust Lagrange multiplier)	E = error problem L = lag problem
0 – 16	7.639	24.462	E / L
17 – 19	8.510	26.330	E / L
20 – 24	8.719	26.002	E / L
25 – 30	8.455	24.942	E / L
31 – 44	5.284	13.347	E / L
45 – 54	8.310	26.009	E / L
55 – 64	8.265	31.011	E / L
65 – 74	5.676	23.523	E / L
75 – 100+	5.820	15.753	E / L

Source: Own calculations

The regression results are as follows:

Cohort 0-16

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.008	0.008	0.008
	(0.005)***	(0.002)***	(0.003)***
emprate	0.000	-0.000	-0.001
	(0.872)	(0.889)	(0.804)
k0_16r	-0.013	-0.006	-0.005
	(0.018)**	(0.267)	(0.344)
URB	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Constant	-0.003	-0.005	-0.004
	(0.242)	(0.037)**	(0.118)
Observations	285	285	285
R-squared	0.179		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 17-19

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.006	0.007	0.007
	(0.039)**	(0.004)***	(0.008)***
emprate	-0.003	-0.002	-0.002
	(0.121)	(0.417)	(0.429)
k17_19r	0.016	-0.001	-0.005
	(0.567)	(0.958)	(0.872)
URB	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Constant	-0.001	-0.004	-0.003
	(0.563)	(0.060)*	(0.178)
Observations	285	285	285
R-squared	0.165		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 20-24

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.005	0.006	0.007
	(0.112)	(0.030)**	(0.029)**
emprate	-0.003	-0.002	-0.002
	(0.168)	(0.363)	(0.358)
k20_24r	-0.004	-0.005	-0.005
	(0.669)	(0.524)	(0.535)
URB	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Constant	-0.000	-0.003	-0.002
	(0.924)	(0.344)	(0.557)
Observations	285	285	285
R-squared	0.164		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 25-30

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.004	0.006	0.006
	(0.252)	(0.045)**	(0.055)*
emprate	-0.003	-0.002	-0.002
	(0.110)	(0.347)	(0.335)
k25_30r	-0.010	-0.005	-0.005
	(0.174)	(0.469)	(0.524)
URB	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Constant	0.001	-0.003	-0.002
	(0.671)	(0.383)	(0.622)
Observations	285	285	285
R-squared	0.168		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 31-44

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.004	0.006	0.005
	(0.094)*	(0.013)**	(0.063)*
emprate	0.001	0.001	0.001
	(0.654)	(0.819)	(0.748)
k31_44r	-0.031	-0.019	-0.022
	(0.000)***	(0.002)***	(0.002)***
URB	-0.000	-0.000	-0.000
	(0.190)	(0.094)*	(0.086)*
Constant	0.003	-0.001	0.001
	(0.277)	(0.547)	(0.829)
Observations	285	285	285
R-squared	0.239		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 45-54

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.005	0.007	0.007
	(0.090)*	(0.008)***	(0.012)**
emprate	-0.002	-0.002	-0.002
	(0.262)	(0.448)	(0.436)
k45_54r	0.015	0.003	0.001
	(0.373)	(0.856)	(0.939)
URB	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Constant	-0.003	-0.005	-0.003
	(0.312)	(0.086)*	(0.232)
Observations	285	285	285
R-squared	0.168		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 55-64

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.007	0.008	0.008
	(0.018)**	(0.003)***	(0.008)***
emprate	-0.003	-0.002	-0.002
	(0.137)	(0.366)	(0.431)
k55_64r	-0.011	-0.005	-0.004
	(0.173)	(0.463)	(0.665)
URB	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Constant	-0.001	-0.004	-0.003
	(0.794)	(0.072)*	(0.173)
Observations	285	285	285
R-squared	0.169		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 65-74

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.004	0.006	0.006
	(0.104)	(0.012)**	(0.021)**
emprate	0.000	-0.000	-0.001
	(0.918)	(0.916)	(0.681)
k65_74r	0.029	0.015	0.014
	(0.001)***	(0.063)*	(0.137)
URB	-0.001	-0.001	-0.001
	(0.020)**	(0.011)**	(0.009)***
Constant	-0.005	-0.006	-0.004
	(0.029)**	(0.007)***	(0.051)*
Observations	285	285	285
R-squared	0.209		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Cohort 75-100+

	(1) erated (OLS)	(2) erated (lag)	(3) erated (error)
small	0.006	0.007	0.006
	(0.027)**	(0.003)***	(0.014)**
emprate	0.000	0.000	0.000
	(0.908)	(0.937)	(0.927)
k75_100r	0.026	0.017	0.018
	(0.000)***	(0.001)***	(0.002)***
URB	-0.000	-0.000	-0.000
	(0.113)	(0.097)*	(0.076)*
Constant	-0.006	-0.007	-0.006
	(0.012)**	(0.002)***	(0.013)**
Observations	285	285	285
R-squared	0.234		

Robust p values in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.?? Summary of the regressions

	Share of small firms	Employm. -rate	Cohort effect	Urbanity (dummy)	R^2 OLS	Moran's I	Spatial lag	Spatial error
0-16	+		(-)	-	0.18	8.05	X	X
17-19	+			-	0.17	8.95	X	X
20-24	+			-	0.17	9.05	X	X
25-30	+			-	0.17	8.75	X	X
31-44	+		-	-	0.24	5.49	X	X
45-54	+			-	0.17	8.63	X	X
55-64	+			-	0.17	8.72	X	X
65-74	+		+	-	0.22	5.81	X	X
75-100	+		+	-	0.24	6.12	X	X

5 Conclusions

More conclusions remain to be written…

- Households (people) in the age where we can assume they have small children have a negative effect on the propensity of firm formation.
- Households (people) that are in the retirement age have a positive effect on the propensity of firm formation.
- Rurality has a positive effect on firm formation.

Our results correspond to other studies in the area, for example Chang, Chrisman and Kellermans (2009) with respect to a positive correlation between firm formation and age.

Appendix

Diagnostic tests for the residuals of the estimated model

If the null hypothesis of no second-order autocorrelation is rejected the GMM estimator could be inconsistent. However, in this draft version of this paper this second order autocorrelation was not tested for. We simply make an assumption that there is no misspecification in the context of autocorrelation of order 2. However, on the other hand spatial autocorrelation is tested for by Moran's I (Greene, 2008). Like autocorrelation in the time domain, spatial autocorrelation means that adjacent observations of the same phenomenon are correlated. Autocorrelation is about proximity in time while spatial autocorrelation is about proximity in two-dimensional space. Thus, spatial autocorrelation is more complex than autocorrelation because the correlation is two-dimensional and bi-directional. In this thesis, spatial autocorrelation is tested for every cross section according to Moran's (1950) I statistic which is computed for each set of the residuals ε_t (or ν_t).

$$I_t = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} W_{i,j}\left(\varepsilon_{i,t}-\bar{\varepsilon}_t\right)\left(\varepsilon_{j,t}-\bar{\varepsilon}_t\right)}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} W_{i,j}\right)\sum_{i=1}^{n}\left(\varepsilon_{i,t}-\bar{\varepsilon}_t\right)^2} \qquad (2)$$

cvbn

I_t is the Moran's I test statistic, ε_t denotes the residuals, and $W_{i,j}$ is the weight matrix in the above formula. The weight matrix measures accessibility or time distances and is specified as the travelling time by car in the Swedish road network (when following the maximum speed limits). The weight matrix consists of estimated time distances between all Swedish municipalities with time distances that are up to 120 minutes. Time distances above 120 minutes are considered as inaccessible in this study, and are therefore disregarded. According to Moran's I statistics there are no significant spatial autocorrelation problems. The estimated coefficients are available in the below table. Since, spatial autocorrelation constitutes no problem; we go on and estimate the panel relationship by the use of Arellano and Bond's technique where spatial autocorrelation is not taken into account.

Moran's I test for spatial autocorrelation

No signs of spatial correlation was found (To be inserted later)				
Moran's I				

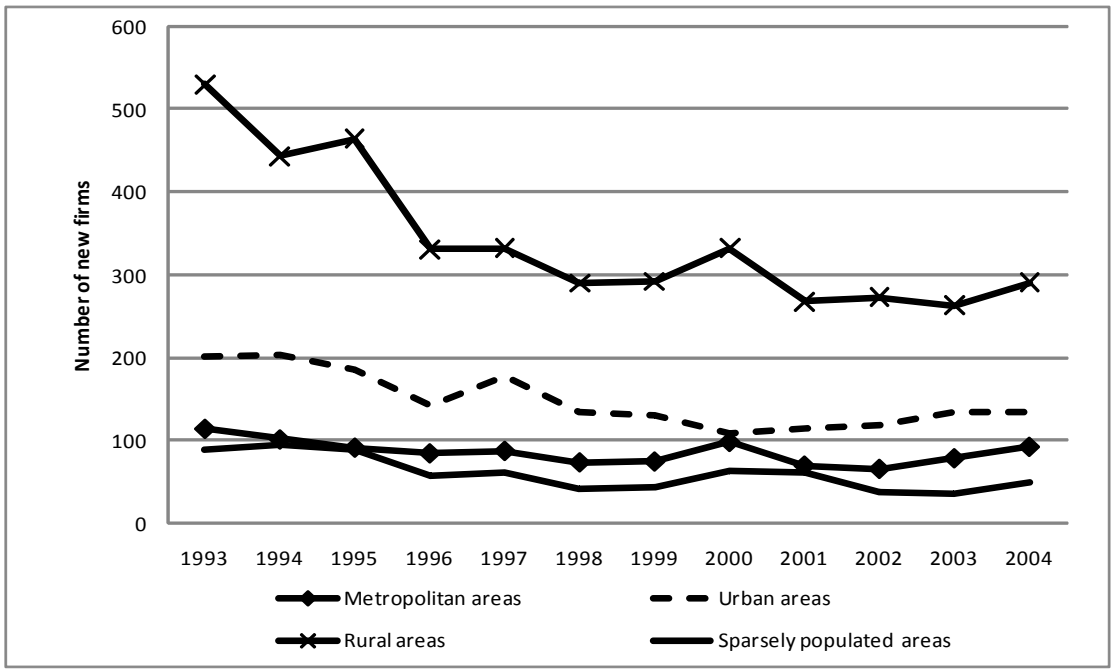

Figure A1: *Number of new firms in rural sectors 1993-2004 in four different types of areas in Sweden*

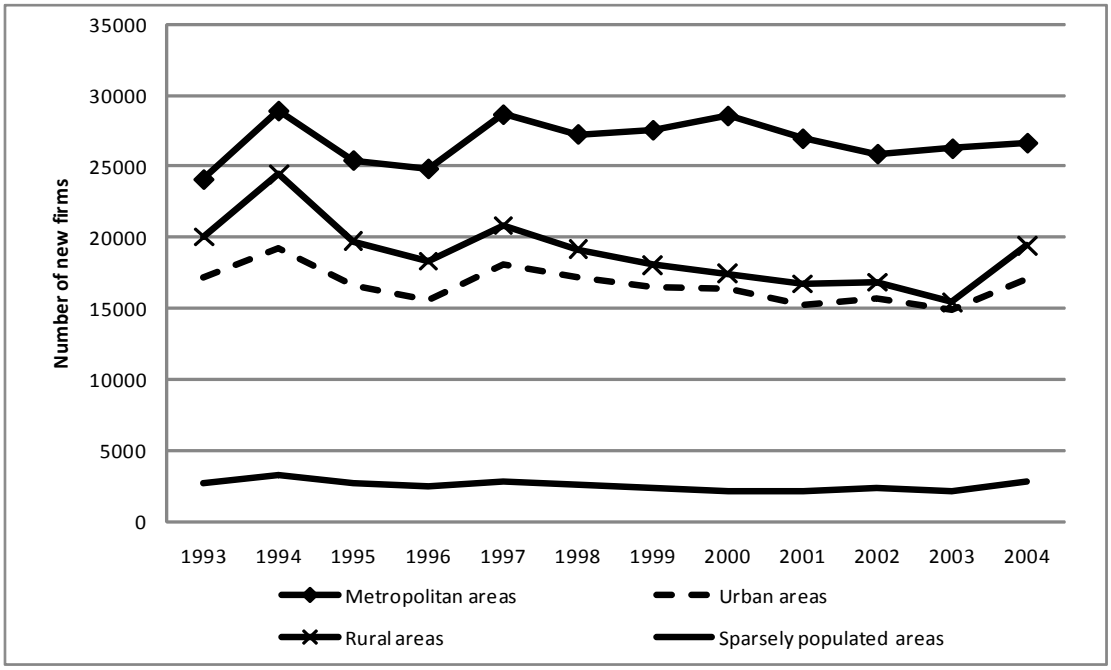

Figure A2: *Number of new firms in other sectors 1993-2004 in four different types of areas in Sweden*

References

Anderson, T. W. and C. Hsiao, (1981), "Estimation of Dynamic Models with Error Components", *Journal of the American Statistical Association*, 76, 598-606.

Arellano, M. and S. Bond, (1991), "Some Tests of Specification of Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 2:277-297.

Amington and Acs (2002) ...

Baltagi, B. H., (2001), *Econometric Analysis of Panel Data,* John Wiley and Sons, New York.

Baumol (1968) ...

Chang, Chrisman and Kellermans (2009) ...

Florida (2002) ...

Florida (2003) ...

Florida (2005) ...

Florida (2008) ...

Fujita, Krugman and Venables (1999) ...

Glesbygdsverket (2007) ...

Glesbygdsverket (2008) ...

Greene, W. H., (2008), *Econometric Analysis,* 6th ed., Prentice Hall, New Jersey.

Holtz-Eakin, D., W. Newey and H. S. Rosen, (1989), "The Revenues-Expenditures Nexus: Evidence from Local Government Data", *International Economic Review,* 30, 2:415-429.

Holtz-Eakin, D., W. Newey and H.S. Rosen, (1988), "Estimating Vector Autoregressions with Panel Data", *Econometrica,* 56, 6:1371-1395.

Hujer, R., P. J. M Rodrigues and C. Zeiss, (2005), "Serial Correlation in Dynamic Panel Data Models with Weakly Exogenous Repressors and Fixed Effects" Working Paper J.W. Goethe-University, Frankfurt.

Johansson and Klaesson (2008) ...

Jordbruksverket (2005) …

Jordbruksverket (2009) RA09:3 …

Kiviet, J. F., (1995), "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models", *Journal of Econometrics,* 68, 1:53-78.

Krugman (1991) ...

Lancaster (1966) ...

Moran, P.A.P. (1950), "Notes on Continuous Stochastic Phenomena," *Biometrika*, 37, 17–33.

Nyström, K., (2006), "Patterns of Entry and Exit in Industrial Sectors in Sweden", in Entry and Exit in Swedish Industrial Sectors, PhD Dissertation, Jönköping International Business School.

Porter (1990) ...

Putnam (1993) ...

Putnam (2000) ...

Quigley (1998) ...

Regeringens skrivelse 2005/06:126 ...

SOU 2006:1 ...

Westlund (2004) ...

And other sources…