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

Conference Paper

Do Startups in the Agricultural Sector Generate Employment in the Rest of the Economy? - An Arellano-Bond Dynamic Panel Study

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)

Do Startups in the Agricultural Sector Generate Employment in the Rest of the Economy? – An Arellano-Bond Dynamic Panel Study

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

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

Lars M Widell
The Swedish Board of Agriculture, Jönköping International Business School and Örebro University, 551 82 Jönköping, Sweden, E-mail: larsm.widell@jorbruksverket.se, Phone: +4636155026.

Abstract

By means of an Arellano-Bond dynamic panel data study of Swedish data over 1993-2004, it is concluded that startups in the agricultural and forestry sector may cause startups in the remaining sectors of the Swedish economy. Thus, new entries in the agricultural and forestry sector may lead to dynamic effects which may lead to employment in completely other sectors of the economy. The agricultural and forestry small-business sector is therefore a very important factor in the quest to reduce unemployment and to increase the economic growth in Sweden. Another important finding, outside of the main purpose of this paper, is that the *per capita* propensity to start a new firm is significantly higher in rural areas compared to urban areas. Consequently, people in rural areas are more entrepreneurial *per capita* (in the context of starting new firms) compared to the population living in urban areas. A substantially high fraction of the rural population faces the forced option of unemployment or self-employment. However, fortunately, a disproportionally high share of the rural population chooses the latter alternative.

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

Startups of new firms are vital elements in the entrepreneurial process to facilitate economic growth, wealth, and employment in a country. Therefore, the point of departure for the purpose of this study is to highlight the obvious fact that *a priori* it is not possible to determine which companies that will be successful over time or get insolvent. Consequently, in this respect, the only feasible option for a nation is to create a business climate that encourages people to start more new firms since a larger population pool of startups will most likely lead to a higher number of companies that in fact will be prosperous in the long run.

The number of startups may be affected by a vast number of factors, but the local population's regional traditions of business activities may be one of the more important factors behind the number of startups. People that are situated in entrepreneurial regions, has contacts, or are in some way related to people with experiences of running a business of their own, are much more likely to pursue a startup. This tendency, in combination with the observation that agriculture is a sector that is dominated by small firms, is a good motive to analyze the question whether startups of firms in the agriculture and forestry sector influence startup activities in other sectors of the economy. Consequently, one of our main purposes with this study is to analyze the old notion that a substantial part of the startups in a country in fact originates from startups in the agricultural sector of the economy.

Therefore, based on an initiative of the Swedish Board of Agriculture, it is analyzed whether startups in the rural sector (ISIC sectors 01-05 with agriculture, forestry, *hunting and fishing*) can explain the variation of startups of firms in the non-rural economic sectors (ISIC sectors 10-93). The main objective of the study is to conclude whether startups in the rural sector may cause multiplicative effects by spreading this growth to the rest of the economic sectors (which henceforth is defined as non-rural sectors). Based on a panel of data over 1993-2004, the main conclusion of this paper is that it in fact can be concluded that startups in the rural sectors seem to cause startups in the remaining sectors of the Swedish economy.

Obviously the stock of firms and the number of startups measure different elements of the economy. However, a historical overview of the Swedish stock of firms reveals that, prior to the 1980s, the number of firms *per capita* decreased at the same time as the largest firms were growing in number of employees. On the other hand, as from the 1980s and onwards there has been an increase in the number of small firms, and this development can, at least partly, be attributed to the growth of the service sector. Consequently, startup of small firms is a vital component in the process to maintain and increase the economic growth in Sweden.

The understanding of how entrepreneurial activities influence economic growth has been important for the design of the Swedish economic policy. The regional policy has

experienced a transformation from a redistributive exogenous perspective to a more endogenous growth perspective. Currently, policy makers are very much aware that on the national level, growth can be considered as the sum of the performance on local and regional level. This shift in how policy is formulated and performed can also be observed on the international level, for example, in policies for growth and development within the EU. An example of this phenomenon is the ongoing transformation of the common agricultural policy (CAP) in Europe, which becomes more and more focused on rural development with policy goals such as stimulation of entrepreneurship, growth, and employment.
In Sweden, a sequentially increasing emphasis on entrepreneurship and how small firms contribute to employment and growth can be observed. A number of analyses have been made with focus on entrepreneurship and growth policy. The startup of new firms does not in itself guarantee that economic growth follows. However, the startup of new firms is by scholars generally considered an important determinant of the long-run development of the economy. In the long-term perspective, it is obviously required that these newly stared firms are fairly successful, profit making, and growing in order to generate significant economic growth for the country. In the literature, we find different arguments for how the startup of firms is stimulated. It is common to split up the arguments for stimulation of startups into push and pull factors. The pull factors are for example good market opportunities and expectations of possibilities of making profits in the future. Push mechanisms are usually exemplified with factors such as unemployment and to traditions of running small firms. In regions with a high share of small firms we are also likely to find a high share of startups of new firms. This also implies that we could expect to find a high share of startups in regions with an industry structure that is characterized by small firms. Small firms are more dominating in a number of industries compared to others. The agricultural industry is a good example where more than 90 percent of all firms do not have any employees at all. Moreover, the agriculture industry hosts almost 10 percent of all firms in Sweden, but at the same time it only hosts around 2 percent of the jobs. Thus, since the agriculture industry is dominated by small firms the focus point in this paper is on whether agricultural firms are able to stimulate startup activities in other sectors of the economy.

Consequently, the main purpose of this study is to analyze whether startups of new firms in the rural sectors (agriculture, forestry, *hunting and fishing*) can explain the variation in startups of firms in the non-rural part of the economy (measured within each respective municipality specification). The empirical analysis conducted by the use of data on local level (municipalities) in Sweden. We utilize a panel-data set of approximately[1] 285 Swedish municipalities for the time period 1993-2004.[2]

[1] Due to political decisions the exact number of total municipalities varies over time.
[2] The estimation period, 1993-2004, may be characterized as a period of recovery in the business cycle directly after the 1992 financial crisis.

2 Entrepreneurship, startups, and job creation

Newly started firms are in general small for obvious reasons. A substantial fraction of all job creation in the economy occurs in small firms. In fact, studies of the US economy illustrates that as much as 90 percent (Allen, 1999) of all job creation originates from small firms. In a study (with similar conclusions) by Davidsson, Lindmark and Olofsson (1994) the authors examined the importance of the small firms in the Swedish economy during the latter part of the 1980s. An interesting finding from the study was that around half of the economic growth and approximately 70 percent of the net creation of new jobs was attributed to small firms during these particular years. Davidsson, Lindmark and Olofsson (1996) followed up their first study but analyzed the recession in the early 1990s, and confirmed the results from their first study. The importance of the small firms was established to be as least as significant during the depression years as during the recovery phase.[3]
Cross country studies indicates that differences between countries with respect to growth performance and job creation can be assumed to be tied to differences in entrepreneurial activities (Global Entrepreneurship Monitor, 1999).

In the economic literature various explanations are given to the reasons for entrepreneurial activities. Entrepreneurship has become a substantial field of research, and an extensive fraction of these studies focus on the question of how startup activities may be explained. Overviews of the doctrine history of how entrepreneurship research has evolved over time can be found in Acs and Audretsch (2003), and it can be concluded that entrepreneurial activities can be explained both from a macro and a micro perspective. In the literature, entrepreneurial activities are found to be related to the overall economic performance, local economic environment, institutional support activities, and individual people's entrepreneurial characteristics.

Although the stock of firms is relatively large in Sweden (if we consider the number of firms *per capita* or *per* employee) the startup activities are on a low level when we make a comparison with other countries (Ekonomifakta, 2009). According to studies of firm formation in Sweden we find different explanations to this observation. A first argument is that in economies with a large number of firms we should expect a somewhat lower level of firm growth compared to countries with a relatively smaller stock of firms because of "catch-up" effects. There are also institutional explanations that depend on the Swedish economy and policy tradition that can be assumed to reduce the propensity to startup firms. Among other factors, these institutional factors are; a high tax rate, a large public sector, and an unbalanced risk distribution for entrepreneurs in Sweden compared to other countries. Thus, these factors may be a part of the explanation why there are relatively few startups in Sweden (Braunerhjelm and Wiklund, 2006). Another reason for the relatively low startup figures may be due to the Swedish industrial structure, where a disproportionally high number of people are working in very large

[3] In the studies by Davidsson, Lindmark and Olofsson (1994, 1996), companies with less than 200 employees are defined as small firms.

firms. Consequently, the stimulation of small-firm startup activities may therefore, to some extent, be constrained by this structural condition.

According to research on entrepreneurship we should expect start up activities of new firms to vary across regions, and to vary over time. Regional economic characteristics such as the size of a region and specialization of the regional industry, etc. can contribute to explain differences in startup activities (Audretsch and Vivarelli, 1996). Geographical variation of entrepreneurial activities is also stressed within the theory of the *new economic geography*. The presence of localized agglomeration economies drives industrial specialization and can also be assumed to have influence on startup activities. External economies of scale that are related to the size of a local/regional industry or the size of a local/regional market stimulate agglomeration of firms. Alfred Marshall (1920) made the notation that firms sometimes can benefit from pooling labor markets and, hence, exhibit increasing returns on a spatial level. Later on, Paul Krugman (1991) and others have developed this type of arguments while also considering how localized skills, spillovers, and other non-tradable specialized inputs can drive location patterns and agglomerations of firms.

In summary, a substantial part of all new jobs are created within small firms. Therefore, startups of more new firms are essential for future economic growth and job creation in Sweden.

3 Urban-rural perspective of the Swedish economy

In this study we do not analyze the behavior on individual or firm level. In the empirical analysis section of this paper, we apply aggregated data on the local level of municipalities in Sweden (there are approximately 285 municipalities in Sweden)[4]. Furthermore, we make a distinction between four types of rural and urban regions in the Swedish economy in order to analyze how these areas differ in various respects, for instance, how the agricultural sector may influence startups in the non-rural part of the economy.

We apply a definition of rural and urban areas that has been applied in numerous other studies of the Swedish economy (Johansson and Klaesson, 2008, SOU 2006:106, and Swedish Board of Agriculture, 2009), where municipalities in Sweden are classified into four different municipality types (MT). These are; (MT 1) metropolitan areas, (MT 2) urban areas, (MT 3) rural areas/countryside, and (MT 4) sparsely populated rural areas. These regional definitions originate from research that has been carried out at the Swedish Board of Agriculture. The four municipality types are defined as follows:

- Metropolitan areas (MT 1): Includes municipalities where the entire population lives in urban agglomerations. In addition to these “core” municipalities this

[4] The exact number of municipalities varies over time. Therefore, an exact number cannot be given in this context.

group includes municipalities where more than 50 percent of the working population commutes to the “core”. Using this definition, there are three metropolitan areas in Sweden; Stockholm, Gothenburg and Malmö, including surrounding municipalities.

- Urban areas (MT 2): Includes 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 which are neighbors to these urban municipalities will be included in a local urban area if more than 50 percent of the labor force in the smaller municipality commutes to a neighbor municipality. In this way, a functional-region perspective is adopted.

- Rural areas/countryside (MT 3): Includes municipalities which 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 kilometer.

- Sparsely populated rural areas (MT 4): Municipalities which are not included in the three categories above and have less than 5 people per square kilometer.

The above categorization of the Swedish municipalities into the four regional subgroups is graphically illustrated in the map below.

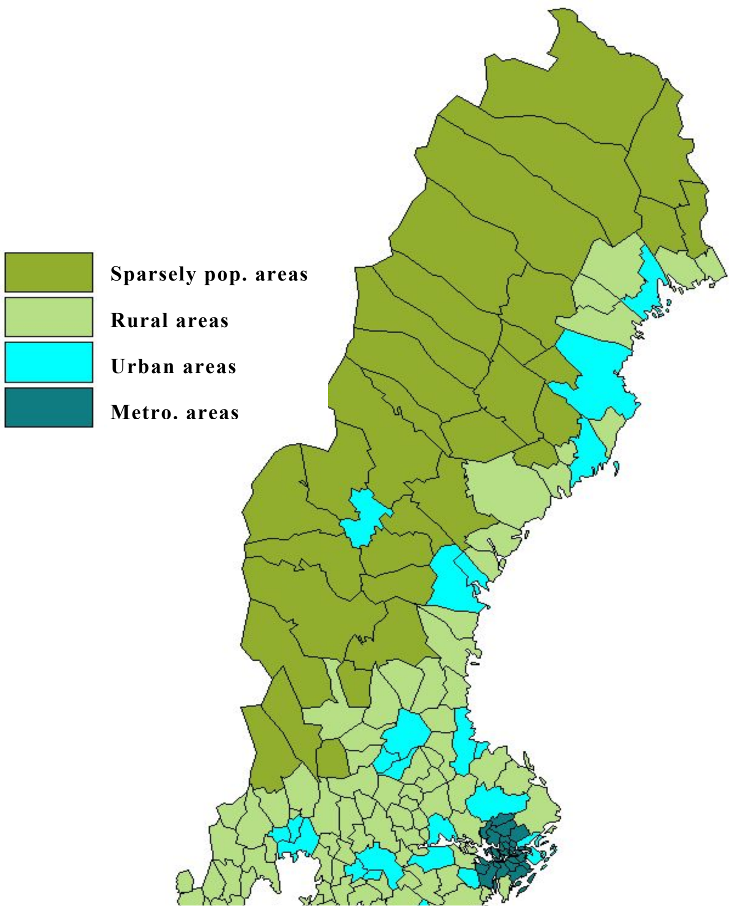

Figure 3.1: *Swedish municipalities classified into (MT 1) metropolitan areas, (MT 2) urban areas, (MT 3) rural areas/countryside, and (MT 4) sparsely populated rural areas*

Despite that we have access to four different municipality types (MT1-MT4) we simplify the discussion in this section by adding together MT1 and MT2 into an urban municipality category, and by adding together MT3 and MT4 which represents the rural part of the economy. This is solely done for pedagogical purposes so that it will be easier to analyze and explain differences between the descriptive statistics for urban and rural areas.

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 decreased their population by 1 percent. Over these years, we can also observe an increase in the number of firms in Sweden by 16 percent. Interestingly, the figures tell us 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 in 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). Consequently, the entrepreneurship is in fact relatively strong in the Swedish 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 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.

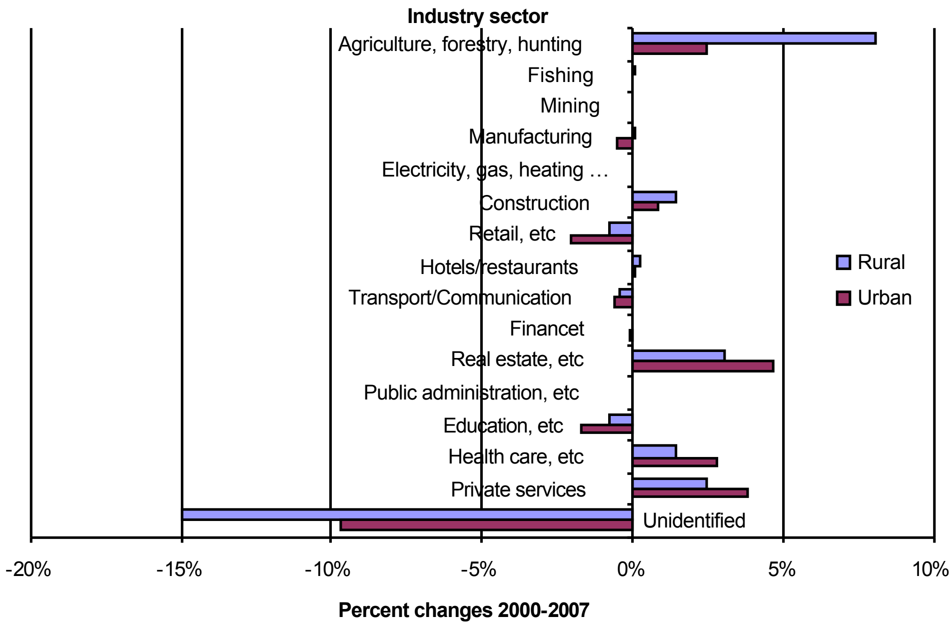

Figure 3.2: *Industry growth 2000-2007 in different industrial sectors in rural and urban areas in Sweden*

Based on an earlier study by the 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 at all. More than 90 percent of the firms in these industries do not have any employees. Table 3.2 below illustrates the structure of work places in rural areas in Sweden in the year 2002. Around 30 percent of all work places in the country were located in rural areas, and around one third of these 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 (2002), Företagsregistret.

Based on stylized facts and on previous research 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 and in various parts of the business cycle.

- 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 is relatively stronger *per capita* in rural areas compared to urban areas. This is because the agricultural sector naturally accounts for a significant proportion of all businesses in rural areas and that this sector is dominated by self employment (small scale operating firms).

- 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 Method, Data, and Empirical Analysis

The main purpose of this study is to analyze whether startups in the rural sector (ISIC sectors 01-05 with agriculture, forestry, *hunting and fishing*) can explain the variation of startups of firms in the non-rural economic sectors (ISIC sectors 10-93). Consequently, the analysis aims to conclude whether startups in the rural sector can cause multiplicative effects by spreading this growth to the rest of the economic sectors (non-rural sectors). This relationship is analyzed by the use of panel data over 12 years (T = 12) and over 285 Swedish municipalities (N = 285) between 1993 and 2004 ($N \cdot T = 3420$ observations). $J_{i,t}$ is defined as the *per capita* number of entering firms in the Swedish rural sectors in municipality i at time t, while the *per capita* entry rate in the non-rural industry sector in municipality i at time t is defined as $E_{i,t}$. Due to the fact that different municipality types are expected to exhibit a large degree of heterogeneity in terms of characteristics and conditions, the data set is categorized into different types of municipality classes. These, previously defined, regional types are (MT 1) metropolitan areas, (MT 2) urban areas, (MT 3) rural areas/countryside, and (MT 4) sparsely populated rural.[5] However, due to the problem of low counts and bottom censoring for MT4, the categories MT3 and MT4 are merged together in a category called rural municipalities (RM = MT3 + MT4).[6] As a consequence, MT1 and MT2 are also merged into a corresponding category defined as urban municipalities (UM = MT1 + MT2).[7] When this type of categorization is conducted there is no risk of bottom censoring. Finally, an universal analysis for all municipalities (AM)[8] is also carried out where we do not split up the data set into any municipality categories. However, a rural dummy variable is included in order to isolate the effect between rural and urban municipalities.

The reason why we categorize the data set into different urban sub-regions is to make an implicit robustness check of our analysis. If the analysis consistently shows the same, or similar, pattern regardless of which municipality categorization that we apply (AM, UM/RM, or MT1-MT4), then we can have much more faith in the result. We can use these categories to analyze whether there are different incentive structures and different patterns of firm formation in rural and urban regions. In addition to the municipality-specific effects, it is also desirable to include time-specific-effects in the model in order to account for the macroeconomic shocks that Sweden experienced during the studied period.

Prior to the discussion regarding the utilized econometric model we present some descriptive statistics over the applied data set. Table 4.1 illustrates count data over the number of new entering firms in the rural sector and in the remaining non-rural sectors of the economy.

[5] MT stands for Municipality Type.
[6] RM is short for Rural Municipalities.
[7] UM represents Urban Municipalities.
[8] AM is an abbreviation for All Municipalities.

Table 4.1: *Descriptive data for entry in rural and non-rural sectors 1993-2004.*

Entry in rural sectors (# firms)	Mean	Std.dev.	Min.	Max.	Skewness	Kurtosis
1993	3.29	3.04	0	15	1.40	5.16
1994	2.96	3.03	0	20	2.10	9.44
1995	2.92	2.90	0	17	1.76	7.36
1996	2.16	2.48	0	14	1.91	7.37
1997	2.31	2.62	0	15	1.96	7.93
1998	1.89	1.99	0	11	1.54	5.65
1999	1.89	2.19	0	14	2.09	9.07
2000	2.12	2.27	0	17	2.33	12.03
2001	1.80	1.78	0	10	1.35	5.22
2002	1.74	1.91	0	12	1.87	8.08
2003	1.79	2.14	0	13	1.97	8.14
2004	1.99	2.07	0	11	1.47	5.28
Entry in non-rural sectors (# firms)	Mean	Std.dev.	Min.	Max.	Skewness	Kurtosis
1993	225	544	20	7954	11.16	150
1994	266	650	22	9735	11.76	164
1995	226	573	20	8500	11.58	159
1996	215	556	15	8192	11.49	156
1997	247	650	23	9725	11.81	165
1998	232	622	24	9224	11.65	160
1999	226	632	12	9487	11.96	168
2000	227	660	16	9891	11.97	167
2001	214	615	15	9265	12.03	169
2002	213	589	15	8760	11.67	161
2003	206	604	13	8982	11.81	163
2004	231	610	22	9184	11.91	168

In order to adjust for non-stationary variables we apply the previously discussed *per capita* transformation of the startup variables ($E_{i,t}$ and $J_{i,t}$). The remaining variables are partly based on a paper by Armington and Acs (2002), but are mainly used as control variables.

The general dynamic panel data model is specified as follows:

$$\begin{aligned} E_{i,t} &= \alpha + \sum_{s=1}^{p} \delta_s E_{i,t-l} + \sum_{u=1}^{q} \beta_u J_{i,t-u} + \sum_{z1=0}^{r1} \theta_{z1} X1_{i,t-z1} + \sum_{z2=0}^{r2} \theta_{z2} X2_{i,t-z2} + \sum_{z3=0}^{r3} \theta_{z3} X3_{i,t-z3} + \sum_{z4=0}^{r4} \theta_{z4} X4_{i,t-z4} + \lambda_t + v_{i,t} \\ J_{i,t} &= \mu + \sum_{u=1}^{q} \pi_u J_{i,t-u} + \sum_{s=1}^{p} \phi_s E_{i,t-s} + \sum_{z1=0}^{r1} \theta_{z1} X1_{i,t-z1} + \sum_{z2=0}^{r2} \theta_{z2} X2_{i,t-z2} + \sum_{z3=0}^{r3} \theta_{z3} X3_{i,t-z3} + \sum_{z4=0}^{r4} \theta_{z4} X4_{i,t-z4} + \gamma_t + \varepsilon_{i,t} \end{aligned} \quad (1)$$

In the above equation (1), δ, β, ϕ, π, and θ are parameters to be estimated by Arellano and Bond's (1991) instrumental-variable GMM technique. The control variables X1-X4 are as follows; educational level (X1) measured as a count variable *per capita*, employment rate (X2), municipality population (X3), and GRP per capita (X4) in each municipality. In order to prevent a high consumption of degrees of freedoms due to the usage of individual municipality specific effects, we estimate the above panel-regression equations four times, once for each municipality type. Consequently, no unobservable individual municipality effects are presented in the above equations, only time-specific effects; γ_t

and λ_t. The error terms $\nu_{i,t}$ and $\varepsilon_{i,t}$ are both assumed to be independently and identically distributed (IID) with zero mean and homoscedastic σ_ν^2 and σ_ε^2 variances. The explanatory variables are assumed to be independent of $\nu_{i,t}$ and $\varepsilon_{i,t}$ for all *i* and *t*.[9]

In the above panel data model, a specific problem arises since the lagged endogenous variable, by specification, is correlated with the disturbance term (even if the disturbance term is not autocorrelated). This problem will cause biased and inconsistent OLS within estimators. Thus, for dynamic panel data models the estimator will be biased but consistent if $T \to \infty$. However, T is fixed in this study which implies that we cannot use OLS since this would cause the within estimator to be inconsistent (Baltagi, 2001).

Moreover it is important to distinguish between dynamic panel data methods and traditional dynamic regression methods, where the latter usually is based on large sample asymptotics ($T \to \infty$) (see e. g. Kiviet, 1995). However, for this paper the dynamic panel data set consists of a small sample in terms of T (T = 12), and in such cases the standard asymptotic methods and the associated tests performed will be seriously biased (Nankervis and Savin, 1987, Kiviet, 1995). OLS and the within estimator will not be utilized in this paper since the results are expected to be biased and inconsistent. As a remedy to this problem, the literature provides several different techniques to solve this difficulty. Since the remedial instrumental variable method suggested by Arellano and Bond (1991) is the most suitable method for our problem, and since it is the one of the most commonly applied methods in applied research (Hujer *et al*., 2005) this method is used in this paper.

Arellano and Bond (1991) demonstrate that additional instrumental variables can be used in the estimation to avoid bias, which makes this method superior to the Anderson and Hsiao (1981) method. The problem is to decide which instrumental variables should be used. We must find instrumental variables that are uncorrelated with the residuals but highly correlated with the original variables. The approach also leads to a relative increase of the efficiency of the estimator. Arellano and Bond (1991) apply a two step GMM estimator which uses the differenced residuals from a GLS one-step consistent estimator. Further details regarding the estimation procedure can be found in Greene (2008) and in Baltagi (2001). Notice, that an important assumption for the Arellano and Bond method of estimating dynamic panel data models is that there is no second-order serial correlation. If there is second-order serial correlation the GMM estimator will not be consistent, however, this problem is empirically tested for in this study.

In accordance with all econometric techniques in the time-series domain, the choice of optimal lag length is very important. There are many different methods for the purpose of determining the optimal lag length. One simple approach is to successively test if the parameters for the last lagged variables are zero. In practice this can be done by a simple Wald test to determine whether the coefficients in the unrestricted model are equal to zero. In this panel data study it is concluded that the optimal lag length is 1 lag.

[9] This assumption is expected to be satisfied since the model is estimated by the Arellano and Bond (1991) method.

However, note that the memory for dynamic parameters has an infinite memory (just as for an autoregressive parameter of order 1 in time-series analysis). The impact of a shock (for an estimated parameter p) will decay by p^t, where t goes from t = 1, 2, 3,... for every period over time. On the other hand, if a non-systematic decay, that is not exponential, would have been found for this process more lags would be necessary. However, in this paper the decay seems to be fairly exponential to its nature which is the reason why 1 lag is assumed to be sufficient. As long as 1 lag represents the correct specification (according to a subjective assessment based on evaluations from model selection criteria, Wald tests, and Arellano and Bond's AR misspecification tests) this single parameter value is easier to interpret since it is often common that the parameters erroneously will obtain varying signs over the lags.

In addition to the instrumental variables, in order to adjust for the time-specific effects time dummies are included in all models. This is specifically important for the Arellano and Bond estimations due to the assumption that there is no existence of second-order serial correlation. Moreover, according to the Moran's I test we could not rule out the occurrence of spatial autocorrelation problems. However, due to the fact that we apply an Arellano-Bond model in this study we cannot adjust the model for possible spatial autocorrelation (additional information regarding this topic is discussed in the appendix).

In Table 4.2 the most important results from this study are presented. The parameters are estimated by the Arellano and Bond (1991) method since the instrumental variables prevents bias and inconsistency due to correlation between the explanatory variables and the residuals.

Table 4.2: *Summary table over selected parameter results from the Arellano-Bond (1991) dynamic panel estimations using instrumental variables*

Dependent variable: E_t	Model for all municipalities (including a rural dummy) (AM)	Model solely for the urban municipalities (UM)	Model solely for the rural municipalities (RM)	Model solely for the metropolitan municipalities (MT 1)	Model solely for the city municipalities (MT 2)	Model solely for the countryside municipalities (MT 3)
$\partial E_t / \partial E_{t-1}$ Influence from startups in E_{t-1}, (δ_1)	0.369^{***}	0.372^{***}	0.176^{***}	0.624^{***}	0.341^{***}	0.232^{***}
$\partial E_t / \partial J_{t-1}$ Influence from startups in J_{t-1}, (β_1)	3.646^{***}	3.022^{***}	4.032^{***}	-0.104	1.085	3.600^{**}
$\partial E_t / \partial J_{t-2}$ Influence from startups in J_{t-2}, (β_2)	2.633^{***}	6.710^{***}	2.399^{***}	0.056	1.159	2.339^{**}
$\partial E_t / \partial J_{t-3}$ Influence from startups in J_{t-3}, (β_3)	4.213^{***}	6.722^{***}	2.400^{***}	2.517	1.300	5.115^{***}
$\partial E_t / \partial J_{t-4}$ Influence from startups in J_{t-4}, (β_4)	5.110^{***}	1.810	1.634^{**}	-0.316	1.166	6.044^{***}
$\partial E_t / \partial X1_{t-z1}$ Influence from the municipality's education rates, $X1_{t-z1}$, (θ_{z1})	0.130^{***}	0.016^{***}	0.370^{***}	0.012^{***}	0.056^{***}	0.222^{***}
$\partial E_t / \partial X2_{t-z2}$ Influence from the municipality's employment rates, $X2_{t-z2}$, (θ_{z2})	0.020^{***}	0.034^{***}	-0.041^{***}	0.019^{***}	0.008^{**}	-0.037^{***}
$\partial E_t / \partial X3_{t-z3}$ Influence from size of the municipality's population rates, $X3_{t-z3}$, (θ_{z3})	7.2e-08***	2.0e-08***	-4.6e-07***	7.9e-09***	1.4e-08	-3.8e-07***
$\partial E_t / \partial X4_{t-z4}$ Influence from size of the municipality's Gross Regional Product, $X4_{t-z4}$, (θ_{z4})	-0.012^{***}	0.004^{***}	0.001	0.001	-0.029^{***}	-0.005^{*}
Rural dummy	0.123^{***}					
AB test for AR(2) in first diff. (p-value)	0.72	0.048	0.058	0.476	0.504	0.360

Dark grey shading = positively sign. parameter point estimate. Light grey shading = negatively sign. parameter point estimate.
* = significant at 10% significance level, ** = significant at 5% significance level, *** = significant at 1% significance level.

The central purpose of this paper is to analyze whether startups in the rural (ISIC 01-05 sector with mainly agriculture and forestry industries) can explain the variation in firm formation in the non-rural economic sectors (ISIC sectors 10-93) for each municipality in Sweden. Therefore, the main empirical finding is that there in fact is an empirical statistically significant relationship between startups in the rural sector and how this is correlated with startups in the non-rural sectors of the economy. In Table 4.2, this is illustrated for the following categories; AM (all municipalities), for UM (urban municipalities) and for RM (rural municipalities). However, interestingly, this relationship can be seen in the countryside (MT3), but not for the city regions (MT2) or for the metropolitan regions (MT1). Thus, the effect in the RM-model is fairly weak.[10] Therefore, the conclusion is that startups of rural (mainly agricultural or forestry) firms may be especially important for the growth of new firm formations in the non-rural economic sectors.

Given that our model is correctly specified we can clearly conclude that firm formation in the rural sector is so important that it leads to significant multiplicative effects in the non-rural sectors of the economy. Consequently, the agricultural and forestry sector is very important for Sweden, and an entry in this rural sector (especially if this entry is in a rural municipality) will lead to multiplicative effects in the remaining sectors of the economy. However, it is very complicated to quantify an exact multiplicator for a new entry in the urban sector, but we can clearly conclude that it is significantly positive.

This is the main conclusion of the study; however, we can also find additional interesting information by interpreting the control variables in the models. For instance, in Table 4.2, the estimated positive coefficients of $X1_{i,t-p}$ tell us that education leads to a higher number of started companies in the non-rural economic sectors. Thus, if people are more educated they seem to be more inclined to start a new firm in the non-rural economic sectors. However, education does not seem to be equally important for the startups of new firms in the rural sector (compared to the non-rural sectors in the economy). This observation makes sense, since education should be expected to be more important in some parts of the service sector compared to e.g. the forestry industry where it may not be necessary with a formal university education to start up a firm.

In Table 4.2, the estimated coefficients for $X2_{i,t-p}$ (employment rate in the studied municipality) tell us that a higher number of new firms are started in urban areas when the employment rate is, or has been, high. However, in rural areas more companies are created when there is unemployment. This may be explained by the fact that this may be the only way to get a new job. Moreover, the government usually induces the possibilities to create new jobs by subsidies and by other measures when the business cycle is down. In urban areas people seem to start new firms when we are in a positive

[10] All these results are confirmed by Granger Causality tests. Moreover, an analogous model as the one in Table 4.2 but where the dependent variable is $J_{i,t}$ (the per capita number of entries in the rural sector) has been estimated but is not presented in this paper. However, the conclusion is that the causality test between $E_{i,t-p}$ in the direction to $J_{i,t}$ (that is, $E_{i,t-p} \rightarrow J_{i,t}$) is weaker than the relationship in the opposite direction (that is, $J_{i,t-p} \rightarrow E_{i,t}$). Thus, even if ($E_{i,t-p} \rightarrow J_{i,t}$) is significant it is clear that the relationship between ($J_{i,t-p} \rightarrow E_{i,t}$) is stronger.

business cycle (when there is a chance to ride on the wave and make a profit), while in rural areas people are more or less forced to choose between unemployment or self employment during bad times. Consequently, in this context, it is not very difficult to explain the different signs of the estimated coefficients in rural and urban areas.

The remaining variables are included as control variables in order to avoid omitted-variable bias. However, there is one last coefficient that is of fairly high interest in this context – the rural dummy in the AM-model. The rural dummy in the model where all municipalities (AM) are included is significantly positive. This tells us that there are significantly more startups *per capita* in rural areas (in comparison to urban areas). Thus, even though the number of startups in absolute terms has decreased lately in the rural geographical regions of Sweden, even more people have migrated from rural areas, which imply that more companies are started *per capita* in rural areas compared to urban areas. Therefore, *per capita*, it is concluded that the rural population is more entrepreneurial than the urban population in Sweden between 1993 and 2004. Despite the evidence in this paper, people will most likely continue to migrate from rural areas to urban areas; however, this established entrepreneurial strength in the Swedish countryside may possibly reduce the intensity of this negative migration process in the future.[11]

5 Conclusions

By means of an Arellano-Bond dynamic panel data study of Swedish data over 1993-2004, it is concluded that startups in the agricultural and forestry *sector* causes startups in the remaining *sectors* of the Swedish economy. Obviously, correlations do not always prove causality, but the correlation relationship is fairly strong in this case and there are many theoretical arguments why the agricultural and forestry sector may cause jobs in other sectors of the economy. Consequently, given that there are no misspecification problems in the model, the agricultural and forestry sector is, due to its dynamic effects, relatively important for the Swedish growth in comparison to many other sectors in the economy. Moreover, according to the above mentioned data set it is established that people are more entrepreneurial *per capita* (in the context of starting new firms) in rural areas compared to urban areas. Rural migration processes and related problems may be explanatory factors of the relatively high entrepreneurship *per capita* in rural areas. A substantially high share of the rural population faces the forced option of unemployment or self-employment. However, fortunately, a disproportionally high share of the rural population chooses the latter alternative.[12]

---

[11] This process of a stagnation of the decrease of migration from rural to urban areas, is also confirmed by figures published by the Swedish Board of Agriculture. Moreover, in Swedish Board of Agriculture (2009) it is also confirmed that the more people *per capita* run a firm in the countryside compared to in urban areas. This is not equivalent to startups, but existing firms may be even more important in this context.

[12] We must bear in mind that on the one hand a large proportion of all start-ups in Sweden can be found in the agricultural sector. On the other hand, this sector is dominated by small firms with low turnovers and limited production volumes. This does not mean that the agricultural and rural businesses are not interesting from a growth

perspective because they are a heterogeneous group of firms producing both goods and services with a multifaceted orientation.

References

Allen K. R., (1999), Launching New Ventures: An Entrepreneurial Approach, Houghton Mifflin, Boston.

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

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

Arellano M. and Bond S., (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.

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

Acs Z. and Audretsch D., (2003), Discoverers of the role of small firms in innovation and creators of the small business economics research field, *International studies in entrepreneurship*, vol. 8, Springer US.

Audretsch D. and Vivarelli M., (1996), Determinants of New-Firm Startups in Italy, *Empirica*, 23(2), 1996, 91-105.

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

Braunerhjelm P. and Wiklund J., (2006), Entreprenörskap och tillväxt – kunskap, kommersialisering och ekonomisk politik, Forum för småföretagsforskning FSF 2006:1.

Davidsson P., Lindmark L. and Olofsson C., (1994), Dynamiken i svenskt näringsliv, Almqvist och Wiksell, Stockholm.

Davidsson P., Lindmark L. and Olofsson C., (1996), Näringslivsdynamik under 90-talet, NUTEK, Stockholm.

Ekonomifakta, (2009), Svenskt företagande internationellt, ekonomifakta, Stockholm.

Global Entrepreneurship Monitor, (1999), 1999 Executive report, Babson College, Kauffman Center for Entrepreneurial Research, London Business School.

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

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

Johansson B. and Klaesson J., (2008), Landsbygdsföretagare måste fixa långa avstånd", Ska hela Sverige leva?, Formas Fokuserar, Forskningsrådet Formas, Stockholm.

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 P., (1991), Geography and trade, MIT Press, London.

Marshall A., (1920), Principle of economics, MacMillan and Co, London.

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

Nankervis J. C. and Savin N. E., (1987), "Finite Sample Distributions of t and F Statistics in an AR(1) Model with an Exogenous Varible", *Econometric Theory,* 3, 387-408.

SOU 2006:106, Fakta – omvärld – inspiration: Underlagsrapporter i arbetet med en strategi för hållbar landsbygdsutveckling, Rapporter till landsbygdskommittén, Stockholm.

Swedish Board of Agriculture, (2005), Åtgärder för konkurrenskraft och tillväxt på landsbygden: bilaga 4 till tekniskt underlag för nytt landsbygdsprogram, Rapport 2005:14c.

Swedish Board of Agriculture, (2009), Företagande på landsbygden: Stad eller land, gör det någon skillnad?, Rapport 2009:2.

Appendix

Moran's I

The central model in this study is presented in Table 4.2. This is an Arellano-Bond model which is estimated by the generalized method of moments (GMM). Therefore, it is by pure definition impossible to estimate a Granger Causality test based on this exact model since R^2 is not defined for this type of model (read more in www.stata.com/statalist). Moreover, we can neither estimate Moran's I (1950) statistic since this model is estimated by GMM. However, in order to approximate whether spatial autocorrelation may be a substantial problem for this data set Moran's I is estimated by the use of a usual OLS cross-sectional model.

Spatial autocorrelation for the estimated model's residuals is tested by Moran's I (1950). Analogously to 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 traditional time-series 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 I (1950) statistic which is computed for each set of the residuals ε_t (or ν_t).[13]

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

I_t denotes Moran's I test statistic, ε_t signifies 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. 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 the estimated Moran's I statistics (in Table A.1 below) there are some tendencies of spatial autocorrelation problems for the earlier years. Since the Moran's I statistic cannot be estimated for the GMM model, the statistics in Table A.1 are only approximations of possible spatial autocorrelation problems. Therefore, it is hard to estimate the potential impact of this possible problem. However, spatial autocorrelation

[13] More information regarding the spatial autocorrelation test by Moran (1950) may be found in Anselin (1988).

does not seem to constitute a very big problem in this study. Therefore, no alternative model to the Arellano-Bond model is estimated.

Table A.1: *Moran's I test for spatial autocorrelation*

Equation / ***Timeperiod***	$j_t = f(j_{t-1}, e_{t-1})$	$e_t = f(j_{t-1}, e_{t-1})$
04=f(03,03)	0.479(0.632)	1.751(0.080)
03=f(02,02)	0.316 (0.752)	0.398(0.691)
02=f(01,01)	-0.202 (1.160)	-0.514 (1.393)
01=f(00,00)	0.444(0.657)	-0.179 (1.142)
00=f(99,99)	0.931(0.352)	0.398(0.691)
99=f(98,98)	0.021(0.984)	1.920(0.055)
98=f(97,97)	0.456(0.648)	0.654(0.513)
97=f(96,96)	2.130(0.033)*	2.910(0.004)*
96=f(95,95)	1.010(0.313)	3.032(0.002)*
95=f(94,94)	2.118(0.034)*	4.676(0.000)*
94=f(93,93)	2.333(0.020)*	0.706(0.480)

Note: p-values within parentheses. * indicates problem with spatial autocorrelation at the 5 % level. The critical value of the test statistic for the 95% confidence level is $Z(I) = \pm 1.96$.