Demographic change, new business formation and spatial concentration of population

Cambiamento demografico, creazione di nuove imprese e concentrazione spaziale della popolazione

Filippa Bono, Maria Francesca Cracolici, Diego Giuliani, Davide Piacentino[1]

Abstract This paper investigates the role of urbanization in the relation between population change and new business formation. The hypothesis of research is that population change significantly influences new business formation only in regions at lower urbanization. This hypothesis is based on the idea that local demand and growing population are significant determinants of new business formation only for rural regions. Using Italian data, we find evidence in support of this hypothesis. Moreover, we find that evidence of spatial dependence in the relation between population change, new business formation and urbanization.

Abstract Il contributo esplora il ruolo dell'urbanizzazione nella relazione tra nascita di nuove imprese e cambiamento della popolazione. L'ipotesi di ricerca è che il cambiamento della popolazione influenza significativamente la nascita di nuove imprese solamente in regioni a più bassa urbanizzazione. Questa ipotesi è basata sull'idea che la domanda locale e la crescita della popolazione sono importanti determinanti della nascita di nuove imprese in regioni rurali. L'analisi condotta sulle province italiana fornisce evidenze a supporto di questa ipotesi di ricerca. I risultati, inoltre, evidenziano la presenza di dipendenza spaziale fra cambiamento demografico, nascita di nuove imprese ed urbanizzazione.

Key words: New business formation; Population change; Urbanization; Spatial models.

---

[1] Filippa Bono, Maria Francesca Cracolici and Davide Piacentino, Department of Economics, Business and Statistics, University of Palermo. V.le delle Scienze Ed.13, 90133 Palermo, Italy. Email: mariafrancesca.cracolici@unipa.it; davide.piacentino@unipa.it
Diego Giuliani, Department of Economics and Management, Università degli Studi di Trento, via Inama 5, 38122 Trento, Italy, e-mail: diego.giuliani@unitn.it.

Introduction

Population decline is generally connected with decline in employment and this may generate a negative spiral where population decline is furtherly caused. As final consequence, a broad decline, from social environment to economics, would influence a region or in general a geographical area. Conversely, growing population would positively influence economics due to increasing demand and business opportunities. In light of these considerations, a positive correlation between population change and new business formation would be expected. This relation is, however, only apparently obvious. Indeed, recent studies pose serious doubts about the obviousness of this relation and, more specifically, they suggest to look at the degree of urbanization (i.e. the spatial concentration of population) as key-factor in this kind of analysis (see Delfmann et al., 2014; Faggio and Silva, 2014). For example, looking at the Dutch case, Delfmann et al. (2014) find that the population change influences new business formation positively in rural regions but negatively in urban regions.

The increasing attention in modern urban economics for the topic of entrepreneurship has reinforced our motivations to provide further evidence on the role of spatial concentration of population in the relation between demographic change and new business formation (see, Glaeser et al., 2010). Our contribution is focused on the Italian case which is particularly interesting due to the significant spatial heterogeneity in terms of both social and economic features.

The paper is organized as follows. Section 2 introduces data and descriptive analysis. Section 3 includes econometric strategy. Section 4 describes empirical results. Section 5 concludes.

Data and variables

This paper mainly exploits data on: (i) demographic characteristics (e.g., demographic dynamics, distribution of population by age, urban density); (ii) social and economic characteristics (e.g., education, self-employment, per capita income, industrial specialization, services, commuting). All variables here employed are based on data collected by ISTAT (Italian National Statistical Office). Table 1 provides a complete list of variables with a short description.

For the sake of brevity, the descriptive analysis is here limited to the rate of new business formation which represents the main variable of interest in this study. We can provide more accurate information on the other variables upon request.

We have measured the new business formation by means of the birth rate of firms (*BR*) as follows:

$$BR = \left[\sum_t \frac{\left(\frac{NF}{LF}\right)}{T}\right] \times 1000 \qquad [1]$$

where, for each year *t*, *NF* is the number of new firms and *LF* is the population in the age-cohort 15-65 (i.e. the labour force).

Table 1: Variables and expected signs

Dependent variables	
BN: Birth rate of firms. Number of new firms for 1000 people. Average 2004-2007. Source: Istat	
Independent variables	***Expected signs***
Pop: Growth rate of population over the time span 2002-2005. Source: Istat	Positive
Urb: Urban density. Population per square km. Source: Istat	Positive
Control variables	***Expected signs***
Age: Difference between the growth rate of each age-cohor and the reference (35-50 age-cohort). Average 2004-2007. Source: Istat	Negative wrt *reference*
Educ: School dropouts rate. Average 2004-2006. Source: Istat	Negative
Income: Per capita income. Average 2002-2005 Source: Istat	Positive
Unempl: Unemployment rate. Average 2002-2005 Source: Istat	Ambiguous
SelfEmpl: Self-employment rate. Average 2002-2005. Source: Istat	Positive
Herfindahl: Specialization rate. Average 2002-2005 Source: Istat	Ambiguous
Services: Specialization in services sector. Average 2002-2005. Source: Istat	Positive
Commuting: Commuting rate. Year 2001. Source: Istat	Positive

Our measure of new business formation is based on the *labour marke*t approach which is grounded on the assumption that each new business is started by individual person (see Audretsch and Fritsh, 1994). Alternatively, *ecological approach* standardizes the number of new firms with respect to the stock of firm in a given region by assuming that new business is generated by incumbent firms (see Van Stel and Suddle, 2008).

Data on firm demography are collected and managed by ISTAT in accordance with procedures suggested by OECD and Eurostat. Data are available for the period 2004-2009. However, due to the fact that the years 2008 and 2009 are strongly

affected by the 2007 financial crisis, we decided to focus the analysis on the sub-period 2004-2007. In other words, we strategically decided to exclude from the analysis an important event which should be cyclical in nature. Data refers to all Italian firms and are aggregated by NUTS3 regions and by sectors of activity.[1]
Figure 1, showing the distribution in quartiles of *BR*, provides evidence of spatial dependence in our data. This evidence is corroborated by Moran's I statistics (I_{Moran} = 0.5058, *p-value* = 0.00). Figure 1 shows that regions with lower BR are located in Southern areas while regions with higher BR are located in the rest of Italy.

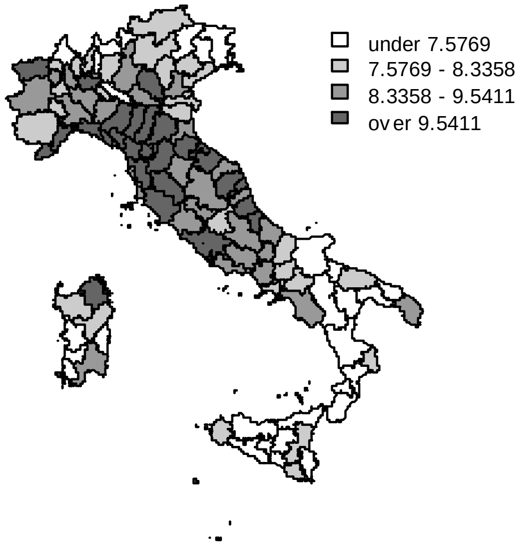

Figure 1: Birth rate of firms

Econometric strategy

The birth rate of firms is characterized by significant spatial autocorrelation, which implies that the regions cannot be treated as independent statistical units. In order to estimate correctly the relationship between birth rate of firms and population change, through a linear regression model, it is therefore necessary to take into account such spatial dependence. Indeed, disregarding the pattern of spatial dependence of data may lead to biased or inefficient estimates of the coefficients (see, for example, LeSage and Pace, 2009).
Spatial econometrics models allows to incorporate explicitly the spatial dependence pattern into the estimation procedure and, as a consequence, to obtain unbiased

[1] The classification of economic activities here used is the NACE Rev 2. Particularly, we aggregate the data into four groups: (i) manufacturing sectors C; (ii) construction sectors F; (iii) trade sectors G, H e I; (iv) other sectors J, K, L, M, N, P, Q, R and S.

results. Amongst the different spatial econometrics model specifications, the *Spatial Durbin Model* (SDM) does not impose a priori restrictions on the order of magnitude of spatial spillover effects (Elhorst, 2010). According to the hypothesis that spillover effects have different intensities for the different explanatory variables that we will include in the empirical model, the proper specification for our aim is the SDM. This model will allow us estimating correctly the direct effect of explanatory variables, on one hand, and also obtaining an estimate of the indirect effect as a measure of spatial spillovers, on the other hand. As we will show in the next section, diagnostic tests confirm the SDM specification. Moreover, LeSage and Pace (2009) have shown that the estimate of model parameters are unbiased even though the true data generation process is a Spatial Lag Model or a Spatial Error Model.

The SDM, including both the spatial lag of the dependent variable, *Wy*, and the spatial lags of the explanatory variables, *WX*, can be represented by the following equation:

$$y = \rho W y + X\beta + WX\theta + \varepsilon \qquad [2]$$

with

$$\varepsilon \sim N(0, \sigma^2 I_n) \qquad [3]$$

where *y* measures the new firm formation rate and *X* is the covariates matrix including both the variable of interest and the control variables. In this equation, the ρ parameter indicates the intensity of the (mean) dependence of new firm formation rates of neighbouring regions and the vector θ includes the parameters of (mean) dependence on spatial lags of covariates.

The β coefficients of the SDM cannot be interpreted ad marginal effects (LeSage and Pace, 2009). Unlike the case of the classical linear regression model, this parameters do not represent the variation induced on variable $\mathbf{y}$ because of a unitary increase of the covariates X. In the SDM specification, a variation of a covariate in the region $\mathbf{i}$ has certainly an effect on the region $\mathbf{i}$ itself, namely the *direct impact*, but also a potential effect, namely *indirect impact*, on the neighbouring regions. LeSage and Pace (2009) have suggested a method to compute properly summary measures of the direct and indirect impacts.

Discussion and concluding remarks

The first model which we estimate is a cross-sectional regression without spatial effects:

$$BR = \alpha + \beta Pop_i + \vartheta Urb_i + \gamma Int_i + \delta Contr_i + \varepsilon \qquad [4]$$

where *BR* is the birth rate of firms; *Pop* is a set of binary variables for four different levels of population change (quartiles are here used); *Urb* is a set binary variables for four levels of urbanization (also in this case, quartiles are used); *Int* is a set of interaction variables between *Pop* and *Urb*; finally, *Contr* is a set of control variables (see Table 1).

We employ the decision rule suggested by Elhorst (2010) for the model selection. Specifically, the OLS model is first estimated and then the robust Lagrange multiplier tests are used to assess which model, spatial lag (LM_{ρ}^{r}) or spatial error (LM_{λ}^{r}) , is more appropriate to describe the spatial structure of data. Table 2 reports, the OLS estimates of two different models (columns 1-2) and the SDM estimates (columns 3-5). The significant values of the robust LM test indicate that a spatial model has to be estimated. Moreover, the LR test indicates that a SDM model has to be preferred.

The SDM estimates do not show a significant direct effect on firm births by population change. It is found a significant and negative indirect effect only for the highest levels of population change (i.e. *Pop3quart* and *Pop4quart*). In other words, this implies that firm births in region *i* negatively depends on population change in other neighbouring regions.

Generally, *Urb* does not exhibit a significant direct and indirect spatial effects. It is found significant and positive effects, both direct and indirect, for the highest level of urbanization (i.e. *Urb4quart*). It is worthy of note that the intensity of the indirect effect is higher than the direct one. An interpretation of such evidence may be found in the high congestion costs which may characterise regions with strong urbanization.

Table 2: Empirical results

Dependent variable: Birth rate of firms (BR)	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
	OLS 1	OLS 2		SDM	
			Direct	Indirect	Total
Population change (Pop)					
Pop1quart	Reference	Reference	Reference	Reference	Reference
Pop2quart	0.079 (0.062)	0.093 (0.036)**	0.061	0.117	0.178
Pop3quart	0.076 (0.071)	0.132 (0.046)***	0.041	-0.317**	-0.277.
Pop4quart	0.161 (0.078)**	0.146 (0.05)***	-0.006	-0.601***	-0.606***
Urban density (Urb)					
Urb1quart	Reference	Reference	Reference	Reference	Reference
Urb2quart	-0.007 (0.059)	0.013 (0.034)	0.024	-0.017	0.007
Urb3quart	-0.068 (0.062)	0.018 (0.038)	-0.048	0.083	0.035

Urb4quart	0.164 (0.085)*	0.088 (0.046)*	0.14**	0.547***	0.687***
Interaction variables					
Pop2quart:Urb2quart	0.019 (0.082)		-0.073	-0.363**	-0.436**
Pop3quart:Urb2quart	0.082 (0.09)		0.13*	0.42**	0.55**
Pop4quart:Urb2quart	-0.04 (0.103)		0.101	0.456**	0.556**
Pop2quart:Urb3quart	0.151 (0.088)*		0.139**	-0.173	-0.034
Pop3quart:Urb3quart	0.164 (0.086)*		0.159**	0.363**	0.522***
Pop4quart:Urb3quart	0.039 (0.1)		0.171**	0.298	0.469
Pop2quart:Urb4quart	-0.134 (0.098)		-0.078	-0.679***	-0.758***
Pop3quart:Urb4quart	-0.029 (0.099)		0.033	-0.15	-0.117
Pop4quart:Urb4quart	-0.097 (0.115)		0.091	0.022	0.114
Control variables					
Age_log(Under15)	0.038 (0.699)	0.208 (0.67)	-0.779	-1.167	-1.946
Age_log(15-25)	-1.802 (0.806)**	-1.166 (0.711)	-0.252	-5.633***	-5.885***
Age_log(25-35)	1.128 (0.759)	0.808 (0.709)	1.562***	2.126	3.688**
Age_log(35-50)	Reference	Reference	Reference	Reference	Reference
Age_log(50-65)	0.123 (0.908)	-0.099 (0.881)	-0.219	4.19**	3.971.
Age_log(Over65)	-3.047 (0.792)***	-2.79 (0.749)***	-2.077***	-5.887***	-7.964***
log(Educ)	-0.015 (0.046)	-0.042 (0.044)	0.02	0.207**	0.226**
log(Income)	0.197 (0.079)*	0.142 (0.073)*	0.2***	0.281	0.481***
log(Unempl)	0.064 (0.06)	0.072 (0.057)	0.032	-0.345***	-0.313.
log(SelfEmpl)	0.193 (0.076)*	0.222 (0.072)**	0.167***	0.293**	0.46***
log(Herfindhal)	0.024 (0.012)*	0.027 (0.012)*	0.036***	0.02	0.056
log(Services)	0.256 (0.124)*	0.246 (0.111)*	0.319***	0.477.	0.796***
log(Commuting)	0.016 (0.021)	0.016 (0.02)	0.054***	-0.002	0.052
R2	0.586	0.578			
LM_{ρ}	33.222***	35.283***	LR test spatial lag		56.639***
LM_{ρ}^{r}	19.998***	8.826***	LR test spatial error		59.304***
LM_{λ}	14.256***	27.338***			
LM_{λ}^{r}	1.032	0.880			

Note: *** 1% significant; ** 5% significant; * 10% significant. Standard errors in brackets.

Interesting results are found looking at the interaction between population change and urbanization. The estimates show significant and positive effects of population change on new business formation in regions with intermediate levels of urbanization. In regions with lower levels of urbanization, we find positive spatial spillovers and their intensity increase with increasing levels of population change. Finally, in regions with higher levels of urbanization, new business formation does not depend significantly by population change. This may be due to the fact that local demand is completely satisfied by the local supply. However new business formation in region *i* seems to be sensible to low levels of population change in its neighbours.

As far as the control variables, the new business formation depends both on demographic characteristics of residents in the region. Regions characterized by a high share of young people are more likely to experiment higher levels of new business formation. Regarding education, new business formation can benefit from a positive spillover effect. The per capita income, self-employment, Herfindahl index (specialization), services sector and commuting significantly influence new business formation.

In conclusion, this study provides evidence in favour of the hypothesis that the relation between demographic change and new business formation is significantly influenced by the degree of spatial concentration of population (namely, urbanization). In particular, our results show that regions with an intermediate degree of urbanization have significant and positive returns in terms of new business formation from a growing population. This relation shows an intensity which increases with increasing levels of population change. On the contrary, we do not find significant evidence for the highest degree of urbanization.

References

1. Audretsch D.B., Fritsh M. (1994). The geography of firm births in Germany. *Regional Studies*, 28(4), 359-365.
2. Delfmann H., Koster S., McCann P., Van Dijk J. (2014). Population change and new firm formation in urban and rural regions. *Regional Studies*, 48(6), 1034-1050.
3. Elhorst J.P. (2010) Applied spatial econometrics: raising the bar, *Spatial Economic Analysis*, 5(1), 9–28.
4. Faggio G., Silva O. (2014). Self-employment and entrepreneurship in urban and rural labour markets. *Journal of Urban Economics*, 84, 67-85.
5. Glaeser E.L., Rosenthal S.S., Strange W.C. (2010). Urban economics and entrepreneurship. *Journal of Urban Economics*, 67, 1-14.
6. LeSage J.P., Pace R.K. (2009) *Introduction to Spatial Econometrics*, Taylor & Francis, Boca Raton.
7. Van Stel A., Suddle K. (2008). The impact of new firm formation on regional development in the Netherlands. *Small Business Economics*, 30(1), 31-47.