Entrepreneurship and unemployment in Spain. A regional analysis*

(Draft March 8th, 2012; please, do not quote)

Begoña Cueto
Matías Mayor
Patricia Suárez

* Department of Applied Economics, University of Oviedo (e-mail: bcueto@uniovi.es)

Abstract:

Self-employment is usually associated to entrepreneurship and it is promoted as a way of reducing unemployment. However, there is no conclusive evidence about this effect of self-employment. Our aim is this paper is to analyse the relationship between self-employment and unemployment taking into account the existence of spatial dependence. The results show small direct and indirect effects. If unemployment increases in a region, self-employment decreases. However, if unemployment grows in neighbouring regions, incentives for entering self-employment raise meaning that there would be a "refugee" effect (self-employment as an answer to the lack of wage employment).

Keywords: self-employment, unemployment, spatial econometrics.

* *Project funded by plan de I+D+I del Ministerio de Ciencia y Tecnología (reference: ECO2009-08061)*

1 Introduction

Entrepreneurship promotion is usually recommended as a way to foster economic and employment growth and innovation. Arguments in favor of policies targeted to entrepreneurs as a way to reduce unemployment are commonly used, especially in economies with high unemployment rates, as the Spanish case.

However, it is quite difficult to identify entrepreneurs. Firstly, we need a definition of 'entrepreneur' but there is no commonly accepted definition. Shumpeter's definition links entrepreneurship to innovation and change in a variety of forms (new goods, new methods of production, new markets, new organization of industries). Other authors have identified entrepreneurship with new opportunities, risk-taking or the value creation (Ahmad and Seymour, 2008).

However, a second issue is how entrepreneurship can be measured in order to have data and information easily available to be considered, especially, in the process of policy proposals. In general, all the definitions include the word 'new' so the amount of entrepreneurs is identified as the number of new business owners. As long as the interpretation of 'new' is stricter, entrepreneurs are measured as the number of new business with employees or the number of high-growth firms (assuming that the novelty is reflected by employment growth). The complexity in the definition makes us to have only proxies for entrepreneurship (Ahmad and Hoffman, 2008). In fact, not only new firms but the older ones can be innovative. So, by extension, the number of entrepreneurs is usually associated to the number of self-employed workers.

The third issue is to have evidence about the relationship of entrepreneurship and unemployment and how self-employment could contribute to economic and employment growth. The evidence is quite extensive for the USA case. Henderson and Weiler (2010) found a strong relationship between entrepreneurial activity (measured by the number of nonfarm proprietors) and economic growth which is greater in more urbanized areas and in the long-run. Results by Rupasingha and Goetz (2011) show that self-employment has a positive effect on economic well-being, income and employment growth, especially in rural communities. Acs and Armington (2004) found a positive relationship between entrepreneurial activity and local economic growth. They found that new firms are more important than the stock of small firms to contribute to economic growth.

For Great Britain, Van Stel and Storey (2004) found no significant relationship between start-ups and employment creation at the national level during the 80s. For the 90s, it is found a positive and significant relationship in the "low enterprise" areas.

Fritsch and Schroeter (2011) analyze the effect of start-up activity on employment for the German case. Their results show that this effect varies across regions. Positive effects are higher in high-density areas, regions with a large share of medium-skilled workers and a high level of innovative activity.

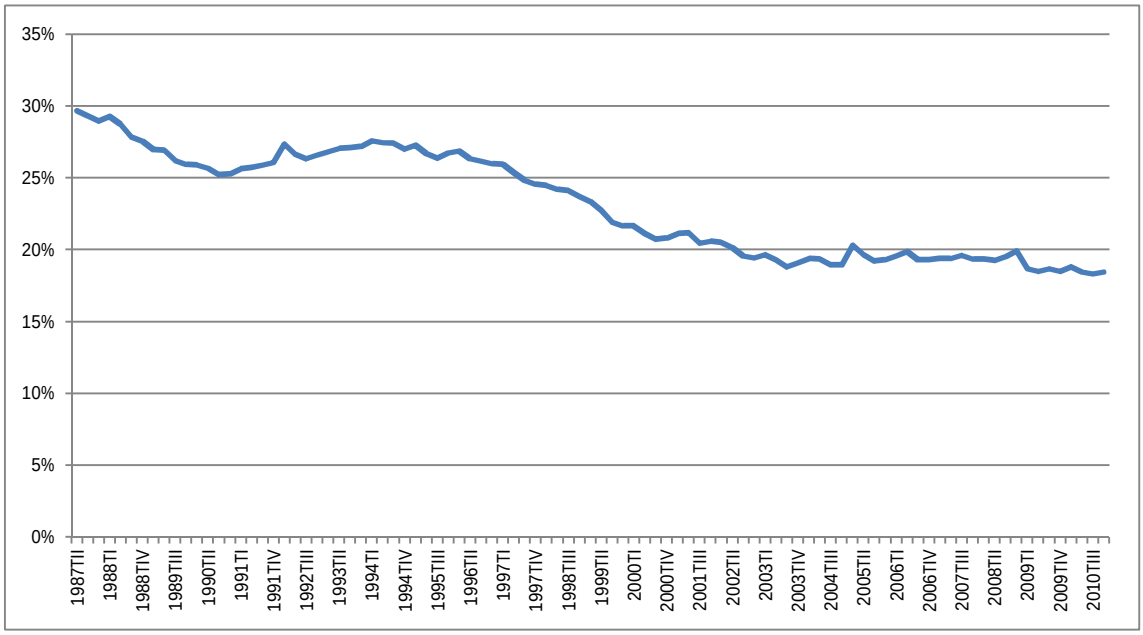
For the Spanish case, Arauzo Carod et al. (2008) found a positive effect of new business formation on employment growth in the manufacturing industries in the short and the long-run. Using a specific sample of the Spanish business structure, García-Tabuenca (2010) found that the 'level of entrepreneurs with creative capability in Spain present values of a low magnitude'. They also found a positive relationship between entrepreneurial activity and the level of economic development.

Summing up, there is no a clear evidence of the positive effect of self-employment on unemployment. Our aim in this paper is to analyse the relationship between these two variables for the Spanish case, taking into account the existence of spatial evidence. In the following section we present the main facts about self-employment in Spain with a special attention to regional differences. After that we summarize the methodology and results are explained. Finally, conclusions and main results are reviewed.

2 Evolution of self-employment in Spain

According to the figures from the Labour Force Survey 3 million people were self-employed workers in 2010 in Spain. If we exclude agricultural sectors, self-employed workers mean 2.6 million over 17.7 million of workers (14.9%). In graph 1 we observe the evolution of the self-employment rate for the period 1987-2010. It is possible to distinguish three periods: an increase from 1989 to 1993, followed by a decrease during 1994-1996 and another phase of stability. The decrease in self-employment rate coincides with the increase in employment since 1994.

Figure 1. Self-employment in Spain, 1986-2010



Source: LFS.

Our analysis refers exclusively to 2000-2008 given that in 2008 a methodological change modifies the definition of self-employment[1]. In Figure 2 we can observe the differences in self-employment rates at the beginning and at the end of our observation period. In general, provinces maintain their positions so self-employment rate is quite stable over time.

Self-employment is not homogeneously distributed across territory. While self-employment rate is below 15 per cent in several provinces, it doubles the national rate in other ones. Madrid has the lowest self-employment rate in the majority of the considered years; Barcelona is also one of the provinces with low self-employment rate. These two regions are identified as ones of the most dynamics of the country in terms of employment or income. Valencia and the Basque Country also have low self-employment rates. Another cluster of low self-employment is located in the South-West of Andalusia. In general, high self-employment rates correspond to inland provinces, with a lower degree of urbanization. In several cases, we can observe that entrepreneurship is high in provinces around centres of development, for instance, provinces around Madrid (except Guadalajara) have high self-employment rates. There

[1] In 2009 there is a methodological change in the definition of self-employment in the Labor Force Survey. New questions allow differentiating self-employed workers from those legally self-employment but wage-workers "de facto".

exists a cluster of self-employment in the West of the country (Extremadura and certain provinces of Castilla-Leon.

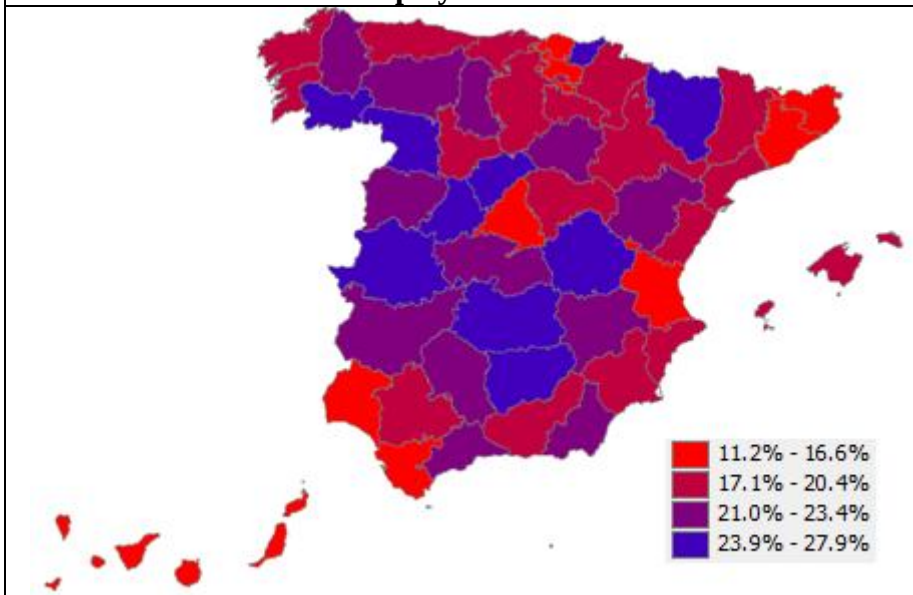
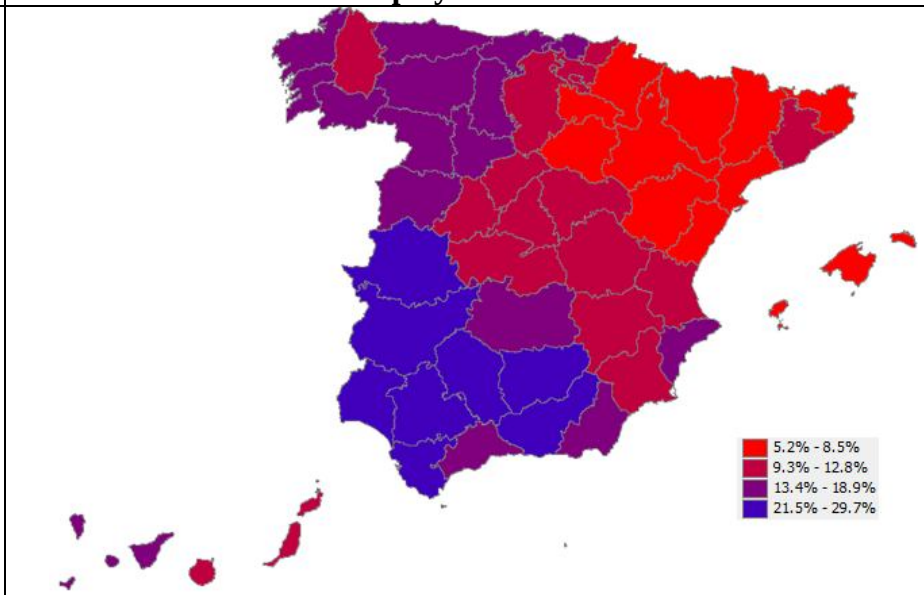
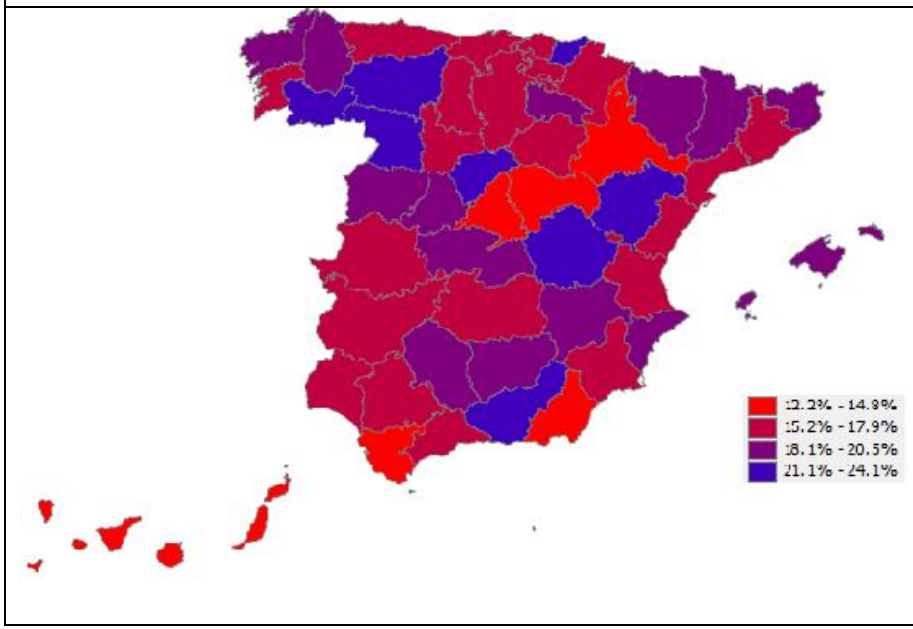
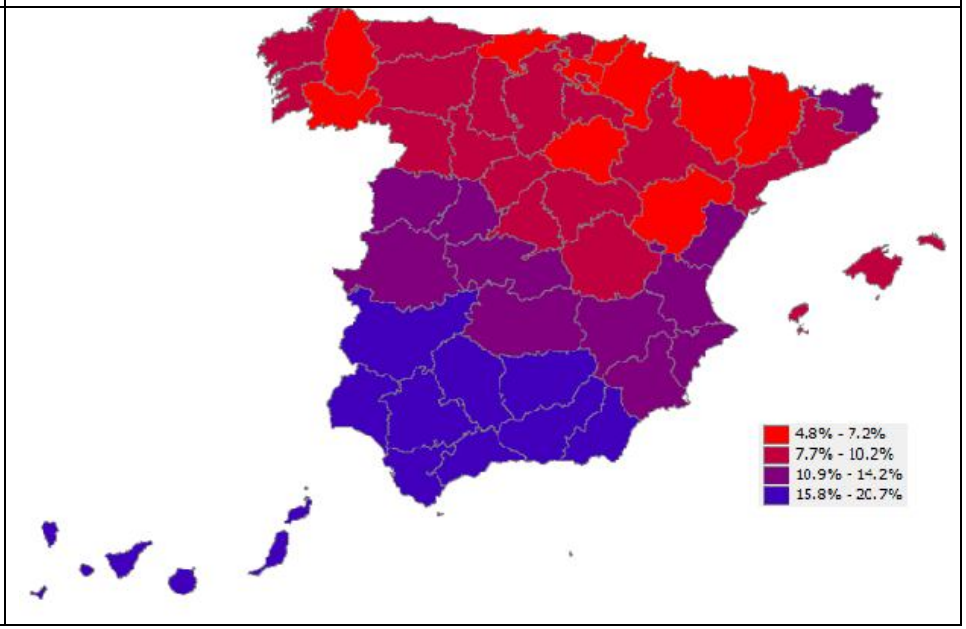
Although there are changes in 2008 (first of all, the self-employment rate is much lower than in 2000), we can keep the previous situation. Madrid is also the province with the lowest level of self-employment and we find the same cluster of high self-employment around Madrid.

Given that our interest is to study the relationship between self-employment and unemployment, in the same figure, we have included maps representing the unemployment rate. This rate is higher in the South than in the North of the country. The lower levels correspond to provinces in the North-East.

In a certain way, we can identify regions with high self-employment rates and unemployment rates (especially provinces in the West). But, at the same time, we also find regions with high self-employment rates and low unemployment like the Basque Country.

In the following section we explain the methodology used in the econometric analysis and results are discussed.

Figure 2. Non-agricultural self-employment rate and unemployment rate

Self-employment rate 2000	Unemployment rate 2000
	
2008	**2008**
	

Source: LFS

3 Methodology and results

As we have seen there are relevant differences among the self-employment rates in the Spanish regions. There are also well-documented differences in the regional unemployment rates. Our interest is to analyse the effect of unemployment on self-employment taking into account the spatial interactions between regions

We have data for the 47 Spanish provinces (excluding Canary Islands, Balearic Islands, Ceuta and Melilla) for the period 2000-2010. The dependent variable in the models is the self-employment rate. Our independent variables are the unemployment rate, the proportion of men and people under 30 in the population and the distribution of the population by level of education.

The unemployment rate is one of the most common variables used in the explanation of self-employment, although few studies have found significant relationships. Depending on the methodology, the data sets and the sample, the relationship between self-employment (rate or probability to enter) and the unemployment rate are different. From a macroeconomic point of view, the recession-push hypothesis suggests a positive relationship between unemployment and self-employment because, during a recession, unemployment acts as a catalyst, encouraging the unemployed to start up in business (Evans and Leighton, 1990; Staber and Bohenhold, 1993). However, a negative correlation is possible. Meager (1992) suggests a second relationship between unemployment and self-employment, labelled the pull hypothesis in the sense that when economic activity levels are growing (unemployment rate falls) more people would enter self-employment because their businesses are less likely to fail.

The demographic composition of the labour force is also included. The results obtained in several microeconomic studies indicate that women and young people are less likely to become self-employed, so the regions with a higher proportion of women and/or people under twenty-five years old would probably have a lower self-employment rate. Evidence on education is mixed: the least educated have high probabilities of being self-employed and there is also evidence that the most highly educated have high probabilities (Blanchflower, 2000).

3.1 Methodology and results

In this paper, we use a panel data structure which extends the modelling possibilities in comparison to the cross-sectional approach applying in previous works. As we described above, the spatial performance of the regional labour markets in Spain plays an important role its consequence, i.e. the existence of spatial dependence may be included explicitly in the modelling framework.

In recent years, there is a growing interest in the development of spatial econometrics techniques for panel data. In this line, Elhorst (2010a, 2012) summarizes the recent contributions on the specification and estimation of dynamic spatial panel data model

At the same time, it is possible to distinguish to different strategies for selecting the spatial econometric model. The first option is known as the specific-to-general approach and consists on testing if there is or not spatial autocorrelation in the OLS residuals obtained from a non-spatial model and propose and spatial or spatial error specification. This strategy is described by Florax et al. (2003) and is based on the Lagrange Multiplier tests and their robust version. This is the most common alternative in empirical works.

The other option is the contrary. The initial model is the most complete option. It nested within it as special cases, a series of simpler models that ideally should represent all the alternative economic hypotheses requiring consideration. Its specification may include the three different types of interaction effects (Mansky, 1993) distinguished when the studied economic variable in one location could be influenced by the behaviour of its neighbouring locations. These effects are (i) the endogenous interaction effects where the decision of a spatial unit depends on the decision taken by other spatial units; (ii) exogenous interaction effects, where the decision of a spatial unit to behave in some way depends on independent explanatory variables of the decision taken by other spatial units and (iii) correlated effects, where similar unobserved environmental characteristics result in similar behaviour. In our case, the endogenous interaction effect collects the idea that the value/ level of the self-employment in one location is determinated not only by the own characteristics but also the values in the neighbouring locations. The interaction effects capture the idea that the economic performance of the spatial units (provinces) is not isolated. Changes in the explanatory variables of self-employment in one location may affect the level of self-employment in

other nearby locations. There is no theoretical support for the third type of effects. This effect collects omitted factor in the model that they are spatially autocorrelated.

In this paper, we really focus on the role of these different types of interaction effects and its estimation more that in the analysis and comparison of the different selection strategies.

The most complete model (called Manski model by Elhorst, 2010) follow this expression with spatial and time-period specific effects:

$$Y_t = \rho W Y_t + \alpha i_N + X_t \beta + W X_t \theta + \mu + \xi_t i_N + u_t$$

$$u_t = \lambda W u_t + \varepsilon_t$$

The spatial and time effects can be considered fixed or random and Hausman's test could be used to test the random effects model against the fixed effects models. In any case, the adoption of random effects may be not adequate when the spatial data covers all provinces. In this situation, Beenstock and Felsenstein (2007), Nerlove and Balestra (2003), among others, point out spatial and/or time fixed effects may be adopted since the values of the variables in each spatial unit is not obtained randomly[2].

Due to the identification problems point out by Manski (1993) if these effects are estimated all together, Lesage and Pace (2009) propose to exclude in these situations the spatial error autocorrelation term. Furthermore, the spatial Durbin model produces unbiased coefficient estimates also if the true data-generation process is any of the other spatial regression specifications except for the Manski model.

Lesage and Pace (2009) modified the common way to analyze the spatial models and test the existence of spatial interactions showing that the direct interpretation of the estimated coefficients are incorrect. Thus, a change in one explanatory variable in the province *i* will not only exert a direct effect on its own self-employment level, but also an indirect effect on the self-employment level of other provinces. Consequently, the interpretation of the effects on dependent variable Y of a unit change in an exogenous variable X_j , the derivative $\partial \mathbf{Y} / \partial \mathbf{X}_j$ is not simply equal to the regression coefficient since it also takes account of includes the spatial interdependencies and simultaneous feedback embodied in the model.

[2] See Elhorst (2012) for a more detail discussion.

As the partial derivative impacts take the form of a matrix $\left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\mathbf{I}\beta_j$, LeSage y Pace (2009) propose new scalar summary measures to collect all these interactions between municipalities so that we may reach a correct interpretation of the spatial models and distinguish between the direct and the indirect impact. The direct impact shows the average response of the dependent variable to independent variables, including feedback influences that arise from impacts passing through neighbors and back to the municipality itself. The indirect impact tackles the effect that any change in a region has on others and how changes in all regions affect a region. Elhorst (2010c) obtained how these effects are calculated when a non-dynamic panel data model is estimated. Since the spatial weight matrix and the parameters β, θ are fixed on time, these are computed in the same way that LeSage and Pace (2009) proposed in a cross-sectional setting. The direct effect is computed as the average of the diagonal elements of the matrix on the right-hand side of (10), and the indirect effect as the average of either the row sums or the column sums of the non-diagonal elements of that matrix.

3.2 Results

We display the results from two models in table 1. The first one includes as independent variables the unemployment rate and the proportion of men and population under 30 in the population. The second one also includes the distribution of the population by level of education. The coefficient δ is significant (0.240 and 0.204) and shows that there is spatial autocorrelation in the Spanish provinces.

According to the coefficients, an increase in unemployment rate in province *i* reduces self-employment in the same province. When we take into account the spatial interactions between provinces this relationship is positive so if unemployment grows in other regions (*j*) self-employment in region *i* increases.

Table 1. Results

	Spatial Durbin I		Spatial Durbin II	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**
Unemployment	-0.071	-2.561	-0.065	-2.354
Men	0.489	1.357	0.657	1.840
Young30	-0.209	-2.648	-0.309	-3.809
Primary	-	-	-0.023	-0.504
University	-	-	-0.042	-0.696
WUnemployment	0.072	2.356	0.051	1.683
WMen	-0.431	-0.584	0.238	0.345
WYoung30	0.638	6.982	0.371	3.093
Primary	-	-	0.071	0.839
University	-	-	-0.213	-2.193
W x dep.var	0.240	4.155	0.204	3.863
Log-L		549.572		557.993

Explaining the magnitude of the coefficients is not correct so we have estimated the direct and indirect effects. LeSage and pace (2009) propose new measures to collect all the interactions between regions so that we may reach a correct interpretation of the spatial models.

The direct effect refers to the average response of the dependent variable to independent variables, including feedback influences from impacts passing through neighbours and back to the region itself. The indirect effect reflects the effect that any change in a region has on others and how changes in all regions affect a region.

In table 2, we display these direct and indirect effects. The direct effect of unemployment is negative while the indirect effect is positive. So an increase in unemployment in region *i* causes a decrease in self-employment in the same region and unemployment growth in other regions (*j*) increase self-employment in region *i*. The magnitude of the effects is significant although small.

Our interpretation of these results is that if unemployment increases in a region, the probability of business survival reduces so self-employment decreases. Even more, the population can try to look for a job in other regions before starting a business. However, if the growth in unemployment is general, incentives for entering self-employment can rise meaning that there would be a “refugee” effect (self-employment as an answer to the lack of wage employment).

When the distribution of the population by level of education is considered in the model, the effects are smaller than in the previous estimation (especially in the case of the indirect effect) so the total effect is negative.

Table 2. Direct and indirect effects

Spatial Durbin Model I	**Direct**		**Indirect**	
	Coeff.	t-stat	Coeff.	t-stat
Unemployment	-0.068	-2.631	0.069	2.165
Men	0.459	1.222	-0.419	-0.430
Young30	-0.176	-2.393	0.741	7.463
Spatial Durbin Model II	**Direct**		**Indirect**	
	Coeff.	t-stat	Coeff.	t-stat
Unemployment	-0.061	-2.312	0.043	1.422
Men	0.704	1.890	0.449	0.546
Young30	-0.294	-3.518	0.368	2.547
Primary	-0.016	-0.351	0.081	0.783
University	-0.055	-0.895	-0.262	-2.321

These results have been estimated for a period of great employment growth so the reduction of unemployment causes an increase in self-employment acting as a push factor. An extension of this research implies to analyse this relationship of the last four year of crisis and employment destruction trying to test if the results are similar.

4 Summary and conclusions

Our aim is this paper is to study entrepreneurship in Spain considering a spatial point of view. Using as a proxy for entrepreneurship the number of self-employed workers, we have explained that there is heterogeneity in self-employment rates in Spain.

It is usual to recommend policies fostering entrepreneurship as a way to increase employment so our interest is to analyze the relationship between self-employment and unemployment using spatial econometrics techniques.

Results show a negative direct effect while the indirect effect is positive. So an increase in unemployment in region *i* causes a decrease in self-employment in the same region and unemployment growth in other regions (*j*) increase self-employment in region *i*.

Our interpretation of these results is that if unemployment increases in a region, self-employment decreases. However, if the growth in unemployment is general, incentives for entering self-employment can raise meaning that there would be a "refugee" effect (self-employment as an answer to the lack of wage employment).

References

Acs, Z.J. and Armington, C. (2004): "Employment growth and entrepreneurial activity in cities." *Regional Studies*, vol. 38(8), pp. 911-927.

Ahmad, N. and Seymour, R.G. (2008): "Defining entrepreneurial activity: definitions supporting frameworks for data collection". OECD Statistics Working Papers, 2008/1, OECD Publishing.

Arauzo Carod, J.M., Liviano Solís, D. and Martín Bofarull, M. (2008): "New business formation and employment growth: some evidence for the Spanish manufacturing industry." *Small Business Economics*, vol. 30, pp. 73-84.

Blanchflower, D.G. (2000): "Self-employment in OECD countries." *Labour Economics*, vol. 7(5), pp. 471-505.

Elhorst J.P. (2010a): Spatial panel data models. In: Fischer MM, Getis A (eds) *Handbook of applied spatial analysis*. Springer, Berlin, Heidelberg and New York, pp 377–407.

Elhorst, J.P. (2010b) Applied spatial econometrics: raising the bar. Spatial Econ Anal 5(1):9–28

Elhorst, J. P. (2012): "Dynamic spatial panels: models, methods, and inferences", *Journal of Geographical Systems*, vol. 14:5-28.

Evans, D. S. and Leighton, L. S. (1989): "Some Empirical Aspects of Entrepreneurship." *American Economic Review*, vol. 79, nº 3, 519-535.

Florax, R. J. G. M., Folmer, H. & Rey, S. J. (2003): "Specification searches in spatial econometrics: the relevance of Hendry's methodology", *Regional Science and Urban Economics*, vol.33, pp.557-579.

Fritsch, M. and Schroeter, A. (2011): "Why does the effect of new business formation differ across regions?" *Small Business Economics*, vol. 36, pp. 383-400.

García-Tabuenca, A., Crespo-Espert, J.L. and Cuadrado-Roura, J.R. (2010): "Entrepreneurship, creative industries and regional dynamics in Spain." *Annals of Regional Science*, vol. 47(3), pp. 659-687.

Henderson, J. and Weiler, S. (2010): "Entrepreneurs and job growth: probing the boundaries of time and space." *Economic Development Quarterly*, vol. 24(1), pp. 23-32.

LeSage, J. and Pace, R.K. (2009): *Introduction to Spatial Econometrics*. CRC Press.

Manski, C. F. (1993): "Identification of endogenous social effects: the reflection problem", *Review of Economic Studies*,6 vol.60, 531-542.

Meager, N. (1992): "Does unemployment lead to self-employment?" *Small Business Economics*, 4, pp. 87-103.

Staber, U. and Bögenhold, D. (1993): "Self-employment: a study of seventeen OECD countries." *Industrial Relations Journal*, vol. 24:2, pp. 126-137.

Rupasingha, A. and Goetz, S.J. (2011): "Self-employment and local economic performance: evidence for US counties." *Papers in Regional Science*, (forthcoming).

Van Stel, A.J. and Storey, D.J. (2004): “The link between firm births and job creation: is there a upas tree effect?” *Regional Studies*, vol. 38(8), pp. 893-909.

Annex

Table A.1. Descriptive statistics

	2000		2005		2010	
	Mean	Sd	Mean	Sd	Mean	Sd
Self-employment	0.204	0.038	0.181	0.028	0.168	0.025
Men	0.491	0.009	0.492	0.011	0.493	0.011
Young30	0.246	0.030	0.223	0.029	0.193	0.025
Unemployment	0.139	0.061	0.092	0.033	0.187	0.058
Primary	0.487	0.071	0.409	0.068	0.360	0.061
Secondary	0.350	0.049	0.391	0.042	0.415	0.036
University	0.163	0.044	0.200	0.052	0.224	0.056
N=47						