XXXVIII CONFERENZA ITALIANA DI SCIENZE REGIONALI

# ASSESSING THE ROLE OF INDUSTRY VARIETY FOR THE CREATION OF INNOVATIVE START-UPS: EVIDENCE FROM ITALIAN LOCAL LABOUR MARKET AREAS

Roberto Antonietti*, Francesca Gambarotto*

"Marco Fanno" Department of Economics and Management

University of Padova

Via del Santo 33, 35123 Padova, Italy

Corresponding author: roberto.antonietti@unipd.it

Abstract

This paper aims at identifying which local factors affect the creation of innovative start-ups in Italian Local Labor Market Areas. Among these factors, the focus is on industry variety and its related and unrelated components. The empirical analysis is based on data on the creation of Italian innovative start-ups between 2012 and 2015 and on the estimation of zero-inflated and hurdle negative binomial models. The estimates show that the innovative start-ups are more frequently created in areas where unrelated variety is higher, whereas no statistically significant effect is found for related variety. Other relevant local attributes are the size of the area, human capital, presence of university, commuting-based self-containment and cultural diversity. Results are also robust to endogeneity.

Keywords: innovative start-ups, entrepreneurial ecosystem, related variety, unrelated variety

JEL classifications: L26, M13, R11

## 1. Introduction

Innovative start-ups are fundamental for innovation systems. They stimulate inventiveness, competition and market dynamics. They implement technologies that increase productivity and explore the consumer needs to produce new products. In doing so, they can be disruptive or simply more efficient in producing better products. However, a large amount of start-ups die within the first three years of activity and one of the main causes is the lack of an appropriate institutional, technological and market 'habitat' (Acs et al. 2016). This means that start-ups are not good *per se* but that, to survive into the market, they require a peculiar entrepreneurial ecosystem (Colombelli et al. 2016; Vivarelli 2004).

Aim of this paper is to investigate which characteristics of the local environment favour the creation of innovative start-ups. The focus is on industry variety and on its related and unrelated components as main vehicles of knowledge spillovers and as elements facilitating the (re)combination of different knowledge sources.

Studies of entrepreneurial ecosystems stress the central role of the local socio-economic conditions on firms' performance. The existent entrepreneurial resources, the local connectors that create entrepreneurial communities and, finally, the entrepreneurial attitude of the ecosystem play a key role in promoting new business. Entrepreneurial ecosystems are places with a thick knowledge base that act as incubators for the breakthrough and re-combination of competencies (Isenberg, 2011; Mason and Brown, 2014). This concept stimulates new investigations on the ingredients that identify and distinguish the most suitable ecosystems for innovative start-ups (Borissenko and Boschma, 2016).

Indeed, a start-up-based economy requires infrastructures that connect, facilitate and promote creativity and knowledge production. On the contrary, European economies are based on ecosystems created to support firms that look for scale economies and standardized production processes leaving a large empty room for creative young firms. This is particularly true for Italy that is suffering a very long economic stagnation with a negative impact on the birth of innovative firms. To overcome this lack and to re-align its position in Europe[1], in 2012 the Italian Government introduced a specific law for innovative start-ups.

A recent empirical analysis shows not only that this policy has been effective in selecting the most innovative start-ups, but also that these latter are different entities with respect to the other start-ups (Finaldi Russo et al., 2016). Comparing the main features of innovative and non- innovative start-ups, the evidence shows that the former are more concentrated in the knowledge-intensive service and high-tech manufacturing sectors, and in the North of Italy. In addition, the major part of them are younger, smaller in size, employ a higher share of intangible assets (like patents or brands) and show a higher propensity to investment. Finally, these firms experience an initial financial debt, and a lower initial turnover, because they do expect a high growth in sales in the future.

The young age, the high propensity to investment in intangible assets and the high growth potential, in front of an initial low economic performance, make this type of organizations as particularly engaged in truly new and cutting-edge projects that require time to be commercialized. The radicalness of their innovations implies the exploration, processing and re-combination of different pieces of knowledge, which is highly localized and context-specific. One key issue then is to assess which type of environment is more capable to support this explorative process. Specifically, whether an environment that facilitates the combination of related knowledge sources or one that stimulates the complementarities among diversified activities.

The role of the business ecosystem on the creation of innovative start-up has been recently investigated by two empirical analyses on the Italian context. Crossing data from the innovative start-ups directory and the Italian NUTS3 geographical areas, Colombelli (2016) studies the role of local knowledge for creative

[1] In 2014, the share of high-growth firms in Italy are 6.8% against the 9.2% of the European Union. They employ 709.769 workers against, for example, the 2.961.954 employed by UK high-growth firms. See: http://ec.europa.eu/eurostat/documents/2995521/7706167/4-26102016-AP-EN.pdf/20f0c515-ed43-45c3-ad6a-ca0b26b36de5

start-ups. The knowledge structure is defined with respect to the local technology structure, i.e. the degree of local technological differentiation, the degree of local complementarity of technology domains and the similarity of them. Results show that innovative newcomers benefit from the local available unexploited technological knowledge. Particularly, both the local related and unrelated varieties of technologies have a positive impact on the generation of innovative start-ups. Other important characteristics of the local environment are the short distance from the main town, the presence of incubators, the quality of the regional workforce and local demand variations.

This paper complements the analysis provided by Colombelli (2016), but it provides two elements of novelty. First, it uses the local industry employment composition, instead of patents, to measure variety. Second, it refers to smaller geographical units than NUTS 3 regions, namely local labour market areas (LLMA). This geographical dimension is more appropriate to study the ecosystem for innovative start-ups because knowledge spillovers tend to decrease with distance (Feldman, 1999) and because LLMA exploit actual labour flows to identify the boundaries of regions instead of mere administrative criteria.

From a very different perspective, Ghio et al. (2016), using the innovative start-ups directory and distributing them on industries and macro-regions, analyse the relationship between university knowledge spillovers and the creation of innovative start-ups. The study underlines the need of proximity between innovative start-ups and university because of the stickiness of academic knowledge. The presence, at the regional level, of individuals with open-mind attitudes also favours the creation of innovative start-ups.

We extend the previous analysis including related and unrelated variety to a series of local factors that characterize the ecosystem for innovative start-ups, including the quality of local human capital and the presence of universities. Our results show that the creation of innovative start-ups is sensitive to higher unrelated variety, rather than to related variety. This means that radical innovation, and highly innovative business creation, stems from places with strongly diversified economies, like large urban areas.

The paper develops as follows. Section 2 reviews three converging strands of literature that investigate the start-up phenomenon. Section 3 presents the datasets used for the empirical analysis, the variables and the estimation strategy. Section 4 discusses the estimation results. Section 5 concludes with policy implication and recommendations for future research.

## 2. Theoretical Background and research hypotheses

Innovation is a social activity that needs a large variety of competencies to recombine ideas, an entrepreneurial attitude to imitate innovators, and a social context that overlaps profit-driven behaviours to social value accumulation (Cooke, 2016; Kirzner, 1997; Tödling et al., 2011).

Different theoretical and empirical analyses have investigated which conditions stimulate the innovative activity. A strand of literature originates from the seminal paper on knowledge spillovers by Glaeser et al. (1992). This concept builds upon the private and public nature of knowledge and has been used to stress the key role of some factors, like: the co-location of firms and universities (Anselin et al., 2000; Audretsch et al., 2004; Bonaccorsi et al., 2014), the effect of local industry specialization or diversification (De Groot et al. 2009), local human capital and entrepreneurship (Acs and Armington, 2004; Glaeser et al., 2010), the concentration of private R&D on innovation diffusion (Audretsch and Feldman, 1996; Wieser, 2005; Hall et al., 2010), the density of economic activities, as a proxy for urbanization economies (Carlino et al., 2007).

The empirical literature originating from Glaeser et al. (1992), however, is inconclusive on whether innovation is more favoured in a highly specialized or in a highly diversified environment (Beaudry and Schiffaureova, 2009; De Groot et al., 2015). This conundrum is partially solved by Frenken et al. (2007), who develop the concept of ‘related variety’. Relying on Jacobs (1969), they argue that innovation is a recombinant process that «necessarily builds on a pre-existing variety of knowledge and artefacts that are

being combined in new ways leading to new products and services» (Content and Frenken, 2016, p. 3). Since the seminal work by Frenken et al. (2007), many studies engaged in distinguishing the different role of related and unrelated variety on economic outcomes. In general, related variety emerges as a driver of employment growth and export diversification, mainly through the creation of new products, and thus, new jobs. Unrelated variety, instead, helps reducing unemployment growth through the diversification of the regional industry portfolio. As pointed out by Content and Frenken (2016), the literature on related variety leaves some questions unanswered. In particular, how radical innovation can originate from unrelated variety. Specifically, the relationship between (related and unrelated) variety and start-up creation has not been properly explored (Boschma, 2016). This paper aims at filling this gap.

Since innovative start-ups are engaged in the development of radical innovations, which (re)combine very different pieces of knowledge, we assume that they are more frequently created where industry variety is higher. Specifically, we assume that such a recombinant activity is easier where unrelated variety is higher[2]. The key role of unrelated variety is also that of preventing job losses after a negative demand shock. Since the projects on which innovative start-ups are involved are very risky, we assume that both the business opportunities and the possibility to minimize losses from unexpected demand shocks are higher in high-unrelated variety contexts.

In order to explain why innovation activities are concentrated in specific areas, like large cities, a second recent strand of literature, based on the entrepreneurial ecosystem approach, focuses on the behaviour of the entrepreneur and his/her local environment. To be generative of innovation, a place must be an interacting environment where a large variety of skills and capabilities are recombined to transform new ideas into entrepreneurial adventures (Stam, 2015; Cook, 2016). It is the proximity of diversity, i.e. the local industry variety, the basis for exploration of new ideas and new markets. In the entrepreneurial ecosystem approach, the focus is on how local communities feed innovation processes through the overlapping of different domains like human capital, markets, culture, support, finance and policy (Stam, 2015, Isenberg, 2011). In this ecosystem, the space *thickness* assumes a generative role for economic development because it produces creativity and economic opportunities. This means that the new value creation depends on the downward and upward causation linkages as well as on the local organization of private and public actors (Stam, 2015, Stam and Bosma, 2015). Even if this approach is potential for explaining the role of the entrepreneur in the innovation process, it is not yet capable to handle the role of related variety in stimulating new economic purposes and recombining the division of labour, especially facing ICT and the digital manufacturing (Frenken and Boschma, 2007; Cook, 2016).

In the managerial literature, conditions that produce the most fertile habitat for innovation are analysed mining the start-up phenomenon. A start-up is a company that grows fast through the transformation of a brilliant idea into a new business and that it accelerates the market fitting thanks to financial investments of venture capitalists (Shane, 2012). Theoretically, the innovative small firm is described in a very different way than the traditional one. It is a temporary organization created to search a repeatable scaling up business model and not a company mainly focused on efficiency goals (Vesper, 1990). In this way, start-uppers are newcomers that contribute to the productivity dynamics transforming - through the creative destruction process - established routines into new businesses (Schumpeter, 1939). However, their survival is not straightforward. In fact, the population of start-ups is very heterogeneous and only few of them are genuinely innovative (Colombelli et al, 2016). Moreover, the selection process triggered by innovative start-ups can produce a negative effect on the labour market reducing the aggregate employment level. Fritsch (2008) underlines that there are several ways through which new businesses can impact positively on employment and regional development. For example, start-ups can amplify markets or create new markets (Audretsch, 1995), generate a greater variety of products and problem solutions, accelerate the structural change

[2] Some examples of Italian innovative start-ups that use different knowledge sources, or that combine elements from very different sectors include firms that produce smart metering systems (nanotechnologies, housing, energy, software and engineering), software applications to virtually manage queues in public offices, firms that use fruit wastes to produce textiles, firms producing drones for the monitoring of vineyard and other plantations.

competing with incumbent firms and stimulate productivity contesting established market positions (Baumol et al., 1988). Moreover, the start-up phenomenon raises some relevant regional effects in terms of employment outcome and economic development (Fritsch, 2008; Fritsch and Weyh, 2006; Fritsch and Mueller, 2004 and 2006). For example, the crowding-out of successful newcomers cannot occur within the same region or the opposite case, newcomers can be locate in other regions and have an impact on the employment level of another region. Finally, it is relevant to remember that the start-up population is very heterogeneous within and between regions because of different technological, market and institutional opportunities (Acs and Audretch. 1990; Acs et al. 2014).

Relying on these strands of literature, we assume that other local characteristics might stimulate the generation of innovative start-ups. Specifically, we assume that, among the others, a key role is played by: (i) the presence of universities; (ii) the human capital of the local workforce; (iii) the density, or thickness, of local economic activities; (iv) labour flows within the region; (v) cultural diversity.

## 3. Empirical analysis

### 3.1. Data

Data on the number of innovative start-ups come from the register of the Italian Chambers of Commerce and the Italian Ministry of Economic Development, specifically from the online directory "innovative start-ups". Innovative start-ups have been created by national Law Decree n. 221/2012 (so called Law Decree "Growth 2.0"). To be considered as an innovative start-up, a firm must fulfil a series of specific requirements. Specifically: a turnover of less than 5 million Euros; being resident in Italy and active from less than 48 months (60 months after Law Decree n. 3/2015); the majority of social capital being owned by physical subjects; being a not-for-profit organization (i.e. no profit redistribution); not being the outcome of a merger of an acquisition; and being focused on the generation and/or commercialization of new products or services with a high technological value. This object further requires the innovative start-up to fulfil at least one of the following additional criteria: employing a significant amount (at least one third) of highly qualified personnel, spend at least 15% (20% before Law Decree n. 76/2013) of its budget in R&D activities, and holding at least one patent, license or an original computer program.

For the purpose of our analysis, the number of innovative start-ups active in Italian local labor market areas (LLMA) between December 2012 and May 2015 is considered. LLMA are identified by the Italian Statistical Institute (Istat) on a functional basis, using workers' commuting patterns as a criterion. The two main advantages of using LLMA are that they are narrower geographical units than NUTS2 and NUTS 3 regions, and so become more useful when trying to measure knowledge spillovers, and that they do not reflect strict administrative borders, but span across different regions and provinces. Using 2011 population census data, Istat identified 611 LLMA.

The sample includes 3,883 innovative start-ups registered between December 2012 and end of May 2015. Figure 1 shows the geographical distribution of innovative start-ups: we see that they are more concentrated in the North of Italy, especially in LLMA belonging to Lombardy and Emilia-Romagna regions, as well as in the largest metropolitan regions, like Milan, Rome, Naples and Turin. Innovative start-ups are, however, present also in Centre and South of Italy: in particular, we observe some clusters in Campania, Apulia and Calabria regions.

FIGURE 1 HERE

Next Section presents the empirical analysis. The main interest is to assess whether local related and unrelated variety can explain this spatial heterogeneity in innovative start-ups localization, once controlled for a series of local confounding factors.

## 3.2. Model and variables

The model we estimate is the following:

$$[1]\ N_i = \beta_0 + \beta_1 RV_i + \beta_2 UV_i + X_i'\beta_3 + \varepsilon_i,$$

where $N$ is the number of innovative start-ups located in LLMA $i$ in the period 2012-2015, $RV$ and $UV$ measure, respectively, related and unrelated variety, while $X$ represents a vector of additional controls observed at the LLMA level. The term $\beta_0$ is the constant, $\beta_{1,}$ $\beta_2$ and $\beta_3$ are the parameters to be estimated, and $\varepsilon$ is the error component, supposed i.i.d.

Related and unrelated variety indicators are taken from Frenken *et al*. (2007). The former measures the weighted sum of the entropy within each two-digit industry of a local market area, and captures knowldge spillovers among firms producing and selling related products and services. The latter, instead, captures the entropy level at the two-digit level and is a measure of industry diversification at the LLMA level. The sum of the two elements gives the entropy at the five-digit level ($VAR$) for each LLMA $i$:

$$[2]\ VAR_i = RV_i + UV_i,$$

$$[3]\ RV_i = \sum_{j=1}^{J} P_j H_j$$

where $P_j$ represents the two-digit employment shares $P_j = \sum_{k \in S_j} p_k$, being $k$ the five-digit industry which falls under the two-digit industry $S_j$ ($j$=1…$J$) and $p_k$ the five-digit employment shares, and:

$$[4]\ H_j = \sum_{k \in S_j} \frac{p_k}{P_k} \log_2 \left( \frac{P_j}{p_k} \right);$$

$$[5]\ UV_i = \sum_{i=1}^{J} P_j \log_2 \left( \frac{1}{P_j} \right).$$

The following variables are included in vector $X$. First, we control for the size of the LLMA using the number of incumbent plants in year 2011 (*# PLANTS*). We prefer to add this variable on the right-hand side of equation 1, instead of using it as the denominator of the dependent variable $N$, to clearly detect the magnitude, and statistical significance, of the size effect on $N$. We do expect more innovative start-ups to be located in larger local market areas, which are characterized by higher local demand and higher presence of local suppliers.

Second, we add a dummy which identifies LLMA as industrial districts (*ID*). These latter are classified by the Italian Statistical Institute (Istat) following a step-by-step algorithm which labels as ID those LLMA with a dominant specialization in manufacturing activities and of a small-medium size. ID are further classified according to the type of manufacturing specialization of the LLMA. Using 2011 Census data, Istat identifies 141 industrial districts. These areas are of particular interest, as they combine a strong manufacturing specialization with a high propensity towards cooperation and vertical disintegration (Brusco, 1982). Moreover, they are not an urban phenomenon; because of the high presence of manufacturng activities, they are mainly located in rural or peripheral areas. Therefore, we do not expect ID to favor the proliferation of innovative start-ups because the ID ecosystem does not provide the necessary diversification conditions for the cross-fertilization of ideas.

Third, we control for the level of human capital of the area. We use two variables: the first is a dummy which takes value 1 if a university is present within the LLMA (*UNIV*), either with the headquarter or with a peripheral unit; the second is the share of employees holding a university degree (*HK*). LLMA with a higher level of human capital should represent, in general, an ideal ecosystem for the generation and proliferation of innovative start-ups. This can occur either through the presence of a university, which acts as an incubator of innovative start-ups and spin-offs or through the presence of a high qualified workforce, which can either choose to create new, innovative activities (perhaps once having registered a patent), or can represent a pool

of specialized labor that can be recruited by innovative entrepreneurs. For these reasons, we do expect both variables to be positively correlated with *N*.

The capability of a local area to generate innovative start-ups may also depend on its degree of trade openness: areas where imports exceed exports, for instance, may suffer employment and business losses because of the competition from foreign countries, whereas areas where exports exceed imports may benefit from new business opportunities. Using ready-to-use information provided by Istat, we define another ordinal variable (*TRADE*) taking the following five values: 0 for LLMA where there is a strong prevalence of imports, 1 where imports moderately exceed exports, 2 where the trade balance is zero, 3 where exports moderately exceed imports and 5 where exports strongly exceed imports. As far as the variable is conceived, we do expect a positive relation between *TRADE* and *N*.

We further include the unemployment rate (*UNEMP*) of the LLMA in 2011. Its impact on *N* cannot be predicted *a priori.* On the one hand, a higher unemployment rate could signal the presence of unexploited labour force, that is a human resource potential that can represent a possible target for regional entrepreneurship policies: in this case, we would expect a positive relationship between unemployment and the number of innovative start-ups. On the other, LLMA with higher unemploymnt rates could represent depressed and deprived areas, where the innovative start-ups would find the most unfavourable places to develop. In this second case, we would expect e negative correlation between *UNEMP* and *N.*

Other two attributes of the ecosystem for innovative start-ups consist in the amount and spatial dimension of the relationships that occur within each LLMA. The former aspect is captured by an index (*FLOWS*) of the relational intensity and is provided by Istat: specifically, it is given by the percentage of (commuting) flows that connect different municipalities within a LLMA (net of those who work and live in the LLMA) on the total amount of possible flows. The index varies between 0 (i.e. the case of a LLMA where nobody commutes across municipalities) and 1 (i.e. the case in which everyone commutes outside the residential municipality): the higher the index, the higher the circulation of people, and knowledge, within a LLMA. For these reasons, this variable can also capture the quality of the local transport system. Therefore, we would expect a positive correlation between *FLOWS* and *N*: a higher number of innovative start-ups is expected to be found in the most dynamic areas, where people move and exchange ideas.

The latter is captured through an index of LLMA self-containment (*SELF*), as provided by Istat. In particular, it corresponds to the minimum value between an index of self-containment from the demand side (*SELF_D*) and one from the supply side (*SELF_S*). The former measures the ratio between the total amount of people who live and work in the LLMA (net of those who work at home, homeless and those who work abroad) on the amount of people who work in the same LLMA (net of those who work at home, of homeless and of those who work abroad). The latter is the ratio between the amount of people who live and work in the LLMA (net of those who work at home, homeless and those who work abroad) and the total amount of people who live in the LLMA (the total amount of people who live and work in the LLMA net of those who work at home, homeless and those who work abroad). *SELF* is equal to the minimum amount of self-containment of the local area: the higher the index, the higher the capability of the local area to be considered a "market", where production, consumption and social activities are spatially concentrated. Such a variable should be positively related to the number of start-ups: a highly self-contained LLMA should concentrate a higher amount of potential market opportunities than a low self-contained LLMA. However, it can also constitute an obstacle to innovative start-up creation where the market results to be too spatially bounded and of limited size.

We also add the population density of the LLMA (*DEN*) through which we capture urbanization economies. As pointed by urban economics literature (Jacobs, 1969; Duranton and Puga, 2001; Carlino et al., 2007), large urban areas favour the birth and growth of innovative activities thanks to their role of incubators during the earliest stages of development, or to the spatial concentration of innovation inputs like R&D laboratories, scientists and financial capital, or to the high chances of cross-fertilization among a variety of different knowledge sources. For all these reasons, we would expect a positive relation between *DEN* and *N*.

To distinguish between cultural and sectoral diversification, we also control for the degree of socio-economic variety of the LLMA. This aspect is captured by the share of foreign citizens (*FOR*) who live in the LLMA: the higher this share, the higher the cultural diversity of the LLMA and so the higher the chances for the creation of innovative activities.

To allow comparison between the average marginal effects, we standardized each continuous variable at zero mean and unit standard deviation. Table 1 shows the summary statistics of all our variables, while Table 2 show their pairwise correlations.

TABLE 1 HERE

TABLE 2 HERE

Finally, we include 19 regional dummies (at NUTS 2 level), to control for region-specific fixed effects related, for instance, to their institutional quality or to the fact of being a target of national or European industrial development policies.

## 3.3. Empirical strategy

When estimating equation 1, two problems arise. First, since $N$ is the discrete, and non-negative, number of innovative start-ups located in each LLMA, we cannot estimate equation 1 through Ordinary Least Squares (OLS). Second, we have 261 LLMAs (42.72% of the sample) with zero innovative start-up in the period of reference. To cope with the first issue, we estimate equation 1 using a count-data model, specifically a negative binomial model, which allows treating the problem of over-dispersion of the data that arises when the variance of the observed distribution of the count variable is larger than the mean. The second issue requires the use of either a zero-inflated (ZINB) or a hurdle (HNB) version of the negative binomial. Both models allow distinguishing among the process that generates the excess of zeros and the process that generates the positive outcomes, but the two differ in the interpretation of the nature of the zeros.

The ZINB assumes that the origins of the zeros are twofold: "sampling" and "structural". While the former means that zeros come by chance, the latter are observed because of a specific structure of the data, and so are not random. In our case, it is like a LLMA that remains randomly without any innovative start-up or a LLMA that, for some specific characteristics, cannot host any innovative start-up. Operationally, the ZINB consists of two parts: a logit (or probit) estimate to predict the excess of zeros and a second separate negative binomial estimate to predict the positive outcomes.

The HNB model, instead, assumes that all zeros are structural, while the origin of the positive outcome is sampling and follows a truncated negative binomial distribution. The HNB requires a logit (or probit) estimate for the probability not to observe a non-zero observation and a separate truncated negative binomial estimate on the positive outcomes.

The choice between the two models is based on traditional information criterion tests, like the AIC and the BIC. However, the two models often produce very similar results, so that, without any strong theoretical justification, the choice rests on reasons of convenience. For our purposes, it is difficult to choose among the two models. Table 3 shows that the performance of the two models is very similar, with the ZINB showing a slightly better performance. In addition, it is difficult to clearly distinguish whether a LLMA randomly or voluntary chose not to have innovative start-ups between 2012 and 2015. For these reasons, we take the ZINB estimates as the reference, and we use those from the HNB as a robustness test (see Appendix, Tables A1 and A2).

TABLE 3 HERE

Another issue is endogeneity. We rely on the fact that the law establishing the innovative start-ups in Italy was promulgated at the end of 2012 represents a sort of policy shock. In this case, reverse causality between $N$ and our measures of local variety should be avoided, or strongly mitigated. However, there could be some

locational attributes, like the local managerial or entrepreneurial attitude, social capital, and financial development[3], that are not observable in the data and can explain both *N* and local variety. In absence of panel data and valid external instruments, we adopt the two-stage approach developed by Lewbel (2012), where identification relies on heteroskedasticity[4].

In short, the model works under two assumptions. First, the error term from the first-stage equation, $\hat{\xi}$, must be heteroskedastic. Second, there should exist a vector of variables *Z*, which can be a subset of the original group of regressors, for which: Cov($Z_i$, $\varepsilon_i * \xi_i$)=0. If these conditions are satisfied, $(Z - \overline{Z})\hat{\xi}$, where $\overline{Z}$ is the mean of *Z* and $\hat{\xi}$ is the estimated error term of the first-stage equation, represents a vector of valid instruments satisfying the usual rank conditions. Therefore, if $\hat{\xi}$ is consistently estimated, it can be used to generate the instruments and the usual tests for weak instrumentation and over-identification can be applied.

## 3. Results

Tables 4, 5 and 6 show the estimation results. Table 4 concerns the ZINB estimates of equation 1 when local variety is measured through the general indicator *VAR*. Table 5, instead, shows the results when local variety is decomposed into *RV* and *UV*. In both tables, the column "Zero Inflation" refers to the logit estimation of the equation that determines whether *N* is zero, while the column "NB" refers to the negative binomial model that is estimated to predict the counts of innovative start-ups for those LLMA for which *N* is strictly positive. The columns "AME" report the average marginal effects for statistically significant variables only. Table 6, instead, presents the results of the IV estimation using the approach developed by Lewbel (2012). Tables A1 and A2 in Appendix show robustness estimates using the HNB method of estimation.

TABLE 4 HERE

From Table 4 we observe that only two variables explain the excess of zeros in *N*, namely *#PLANTS* and *SELF*: no innovative start-ups are more frequent in small and highly self-contained LLMA. The marginal effect of the former is, however, much larger. Instead, a higher *N* is related to: a larger LLMA size, the presence of a university, a higher local human capital stock, a higher rate of unemployment, a more intense frequency of commuting flows, a higher degree of self-containment of the LLMA, a higher population density and a higher cultural diversity. In line with our expectations, a higher *N* is also related to a higher local sectoral variety, all else remaining equal[5]. Looking at the AME, we observe that the highest impacts concern size, local human capital and variety: in particular, we estimate that a unit increase in *VAR* is related to 3.8 additional innovative start-ups. Local skill availability, an easier and more intense local circulation of people, and a highly diversified local ecosystem are all factors favoring the proliferation of innovative start-ups. On the contrary, being an industrial district is related to a lower *N*: highly specialized and rural LLMA, like industrial districts, constitute a kind of hostile environment for the development of innovative start-ups.

TABLE 5 HERE

Table 5 confirms our expectations: the only element of *VAR* that explains a higher *N* is *UV*, while the estimated coefficient of *RV* is not statistically significant. A unit increase in unrelated variety is related to 3.7

---

[3] Unfortunately, data on entrepreneurship rates, social capital and the banking system are not available at the LLMA level. On the interplay between banks and variety at the local level, see Antonietti et al. (2014).

[4] For recent applications of the Lewbel's approach see, among others: Emran and Shilpi (2012), Millimet and Roy (2015).

[5] The highly statistical significance of $\alpha$ confirms the presence of over-dispersion and that a negative binomial model better fits the data than a Poisson. In addition, the Vuong test indicates that a zero-inflated model is preferred to a standard negative binomial.

additional innovative start-ups, all the rest remaining equal. In line with our hypothesis, the creation and location of an innovative start-up depends on the possibility to combine a wide set of local different knowledge sources.

Finally, Table 6 presents the IV estimation results, using the Lewbel (2012) approach. Column 1 refers to the estimation results on the entire sample of 611 LLMA, while Column 2 show those related to the reduced sample of 350 LLMA with a strictly positive amount of innovative start-ups. In both cases, the Breusch-Pagan test on the first-stage equation strongly rejects the null hypothesis of homoscedasticity in the error component, confirming that using the Lewbel (2012) approach was a right choice. The Hansen J tests also never reject the null hypothesis of correct specification of the models, thus excluding over-identification. The Kleibergen-Paap (KP) F statistics are above the value of 10, but lie between the 5% and 10% Stock and Yogo weak identification critical values[6]. The generated instruments, thus, can be considered as moderately strong. In both columns, the estimated coefficient remains positive and statistically significant, while their magnitude is lower than that resulting from ZINB estimates in Table 5: on average, a unit increase in *UV* generates 2.1-2.2 additional innovative start-ups.

TABLE 6 HERE

## 5. Conclusions

In this paper we investigate the phenomenon of innovative start-ups looking for the environment features that facilitate and support their creation in Italian local labour market areas. Our estimates show that higher unrelated variety significantly affects the amount of innovative startups. This occurs because, in a very diversified local economy, young and innovative firms are more capable to explore the re-combination of local capabilities in new economic pathways. Interestingly, industrial districts do not represent a suitable ecosystem for the proliferation of innovative start-ups.

Other local factors are also relevant, like the size of the local market area and the relational intensity that facilitates knowledge flows among workers, the cultural diversity of local population and the thickness of the local economy, captured through employment density. The presence of a university as well as the local endowment of human capital are key local drivers of innovative start-up creation.

We contribute to the extant literature in different ways. First, we show that unrelated variety matters for the creation of innovation, especially that with a radical content. While incremental innovation, or the innovative activity of incumbent firms, may benefit from related variety, radical innovation and new business engaged in high-risk projects require a different local scenario, where the unrelated variety offers better chances for the cross-fertilization of ideas and business opportunities. This result partially fills a gap in the related variety literature, which tends to over-emphasize the role of related variety.

Second, we expand the literature on entrepreneurial ecosystems, showing the mix of factors that all together depict the place where innovative activities are more stimulated. These factors do not include only industry-related characteristics, but also social and cultural elements, like the cultural openness of local population and labour mobility. Radical ideas tend to rise where socio-economic diversity is higher and where people, and ideas, move more frequently.

These results stress the crucial role of public agency in creating the right infrastructures to support the creation of new businesses. To support innovative start-ups, policymakers have to develop soft infrastructures that support university-industry linkages, local interaction for inventors, skilled people and entrepreneurs, local animation for networking and discovering. This means that, the Italian Law for innovative start-ups is not sufficient as institutional change to trigger the virtuous cycle of the new digital era. To support the catching-up process, Italian policymakers have to map territories and their features in order to define a new growth strategy, i.e. to increase employment and to re-lunch regional economies.

---

[6] The use of Kleibergen-Paap F statistics, instead of the standard Cragg-Donald F statistics, is due to the clusterization of the standard errors.

## References


Acs, Z.J., Armington, C. (2004). The impact of geographic differences in human capital on service firm formation rates, *Journal of Urban Economics*, 56(2): 244-278.

Acs, Z. J., & Audretch, D. B. (1990). *Innovation and small firms*. Cambridge: Cambridge University Press.

Acs, Z. J., Audretch, D. B., Lehmann E.E., Licht G., (2016). National systems of entrepreneurship, *Small Business Economics*, 46:527-535.

Acs, Z.J., Autio, E., Szerb L. (2014) National Systems of Entrepreneurship: Measurement Issues and Policy Implications, *Research Policy*, 43(3):476-449.

Anselin, L. Varga, A., Acs, Z.J. (1997). Geographic and sectoral characteristics of academic lnowledge externalities. *Papers in Regional Science*, 79(4), 435-443.

Antonietti R., Cainelli G., Ferrari M. and Tomasini S. (2015), Banks, related variety and firms' investments, *Letters in Spatial and Resource Sciences*, 8(1), 89-99.

Audretsch, D. B. (1995). *Innovation and industry evolution*. Cambridge: MIT Press.

Audretsch, D. B., Feldman M.P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, 96(3), 630-640.

Audretsch, D., Lehmann, E., Warming, S. (2004). University spillovers: Does the kind of science matter?, *Industry and Innovation*, 11(3): 193-206.

Baumol, W. J., Panzar, J. C, & Willig, R. D. (1988). Contestable markets and the theory of industry structure, revised edition. San Diego: Harcourt Brace Jovanovich.

Beaudry, C., Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* 38: 318–337.

Bonaccorsi, A., Colombo, M. G., Guerini, M. Rossi-Lamastra, C. (2014). The impact of local and external university knowledge on the creation of knowledge-intesive firms: Evidence from the Italian case, *Small Business Economics*, 43(29, 261-287.

Borissenko, Y., Boschma, R. (2016) A critical review of entrepreneurial ecosystems: towards a future research agenda, No 1630, *Papers in Evolutionary Economic Geography* (PEEG), Utrecht University, Section of Economic Geography.

Boschma, R. (2016). Relatedness as driver of regional diversification: a research agenda, *Regional Studies*, online version, http://dx.doi.org/10.1080/00343404.2016.1254767.

Breschi S., F. Lissoni (2001), Knowledge Spillovers and Local Innovation Systems: A critical Survey, *Industrial and Corporate Change*, 10(4): 975-1005.

Brusco, S. (1982), The Emilian model: productive decentralization and social integration, *Cambridge Journal of Economics*, 6(2), 235-261.

Carlino G.A., Chatterjee S and Hunt R.M. (2007), Urban density and the rate of invention, *Journal of Urban Economics*, 61(3), 389-419.

Colombelli A. (2016), The impact of local knowledge bases on the creation of innovative start-ups in Italy, *Small Business Economics*, 47, 383-396.

Colombelli A., Krafft, J., Vivarelli, M. (2016). To be born is not enough: the key role of innovative start-ups, *Small Business Economics*, 47: 277-291.

Content, J., Frenken, K. (2016). Related variety and economic development: a literature review, *Papers in Evolutionary Economic Geography* n. 16.21, Utrecht University.

Cooke, P. (2007). To Costruct Regional Advantage from Innovation Systems First Build Policy Platforms, *European Planning Studies*, 15(2): 179-194.

Cooke, P. (2016). The virtues of variety in regional innovation systems and entrepreneurial ecosystems, *Journal of Open Innovation: Technology, Market and Complexity*, 2:3-19.

De Groot, H.L.F., Poot, J., Smit M.J. (2009) Agglomeration Externalities, Innovation and Regional Growth: Theoretical Perspectives and Meta-Analysis, in Capello, R., Naijkamp, P., *Handbook of Regional Growth and Development Theories*, Edward Elgar, London.

Duranton G. and Puga D. (2001), Nursery cities: urban diversity, process innovation, and the life cycle of products, *American Economic Review*, 91(5), 1454-1477.

Emran M.S. and Shilpi F. (2012), The extent of the market and stages of agricultural specialization, *Canadian Journal of Economics*, 45(3), 1125-1153.

Feldman, M.P. (1999) The New Economics Of Innovation, Spillovers And Agglomeration: A review Of Empirical Studies, *Economics of Innovation and New Technology*, 8:1-2, 5-25

Finaldi Russo, P., S. Magri, C.Rampazzi (2016). Innovative start-ups in Italy: their special features and the effects of the 2012 law, *Banca d'Italia, Occasional Papers*, n. 339.

Frenken K., van Oort F. and Verburg T. (2007), Related variety, unrelated variety and regional economic growth, *Regional Studies*, 418(5), 685-697.

Frenken K., Boschma, R. (2007). A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process, *Journal of Economic Geography*, 7(5): 635-649.

Fritsch, M. (2008). How does new business formation affect regional development? Introduction to the special issue, *Small Business Economics*, 30: 1-14.

Fritsch, M., Weyh, A. (2006). How Large are the Direct Employment Effects of New Businesses? An Empirical Investigation for West Germany, *Small Business Economics*, 27(2): 245-260.

Fritsch, M., Mueller, P. (2004). The effects of new business formation on regional development over time. *Regional Studies*, 38, 961–975.

Fritsch, M., Mueller, P. (2006). The evolution of regional entrepreneurship and growth regimes. In M. Fritsch & J. Schmude (Eds.), *Entrepreneurship in the region*. New York: Springer.

Ghio, N., Guerini, M. & Rossi-Lamastra, C. (2016). University knowledge and the creation of innovative start-ups: an analysis of the Italian case, *Small Business Economics*, 47: 293.

Glaeser, E. L., H. D. Kallal, J. S. Scheinknam and A. Shleifer (1992). Growth in Cities, *Journal of Political Economy*, 100: 1126–1152.

Glaeser, E.L., Rosenthal, S.S., Strange, W.C. (2010). Urban Economics and Entrepreneurship, in Special Issue: Cities and Entrepreneurship Sponserd by the Ewing Marion Kauffman Foundation, *Journal of Urban Economics*, 67(1): 1-14.

Hall, B.H., Mairesse, J. & Mohnen, P. (2010). Measuring the Returns to R&D. In Hall, B.H. and Rosenberg, N. (Eds*). Handbook of the Economics of Innovation*. Amsterdam: Elsevier.

Isemberg, D.J. (2011). The Entrepreneurship Ecosystem Strategy as a New Paradigm for Economic Policy: Principles for Cultivating Entrepreneurship, The Babson Entrepreneurship Ecosystem Project, Babson College, Massachussetts.

Jacobs J. (1969). *The Economy of cities*, New York, Random House.

Kirzner, I. (1997). Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach, *Journal of Economic Literature*, 35(1): 60-85.

Lewbel A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogeneous regressor models, *Journal of Business & Economic Statistics*, 30(1), 67-80.

Mason C., Brown R. (2014). Entrepreneurial Ecosystems and Growth Oriented Entrepreneurship, Background paper prepared for the workshop organised by the OECD LEED Programme and the Dutch Ministry of Economic Affairs on Entrepreneurial Ecosystems and Growth Oriented Entrepreneurship The Hague, Netherlands.

Millimet D.L. and Roy A.J. (2016). Empirical tests of the pollution haven hypothesis when environmental regulation is endogenous, *Journal of Applied Econometrics*, 31(4), 652-677.

Schumpeter, J. A. (1939). *Business cycles: A theoretical, historical and statistical analysis of the capitalist process*. New York: McGraw-Hill.

Shane, S. (2012). Reflections on the 2010 AMR decade Award: Delivering on the promise of entrepreneurship as a field of research. *Academy of Management Review*, 37(1), 10–20.

Stam, E. (2015). Entrepreneurial Ecosystems and regional policy: A Sympathetic Critique, *European Planning Studies*, 23(9): 1759-1769.

Stam, E., Bosma, N. (2015). Local Policies for high-growth firms, in Audretsch, D., Link, A.N., Walshok, M.L. (eds), The Oxford Handbook of Local Competitiveness, Oxford: Oxford Un. Press.

Tödling F., P. Prud'homme van Reine, S. Dörhöfer (2011). Open innovation and regional culture – Findinds from different industrial and regional settings, *European Planning Studies*, 19(11): 1885-1907.

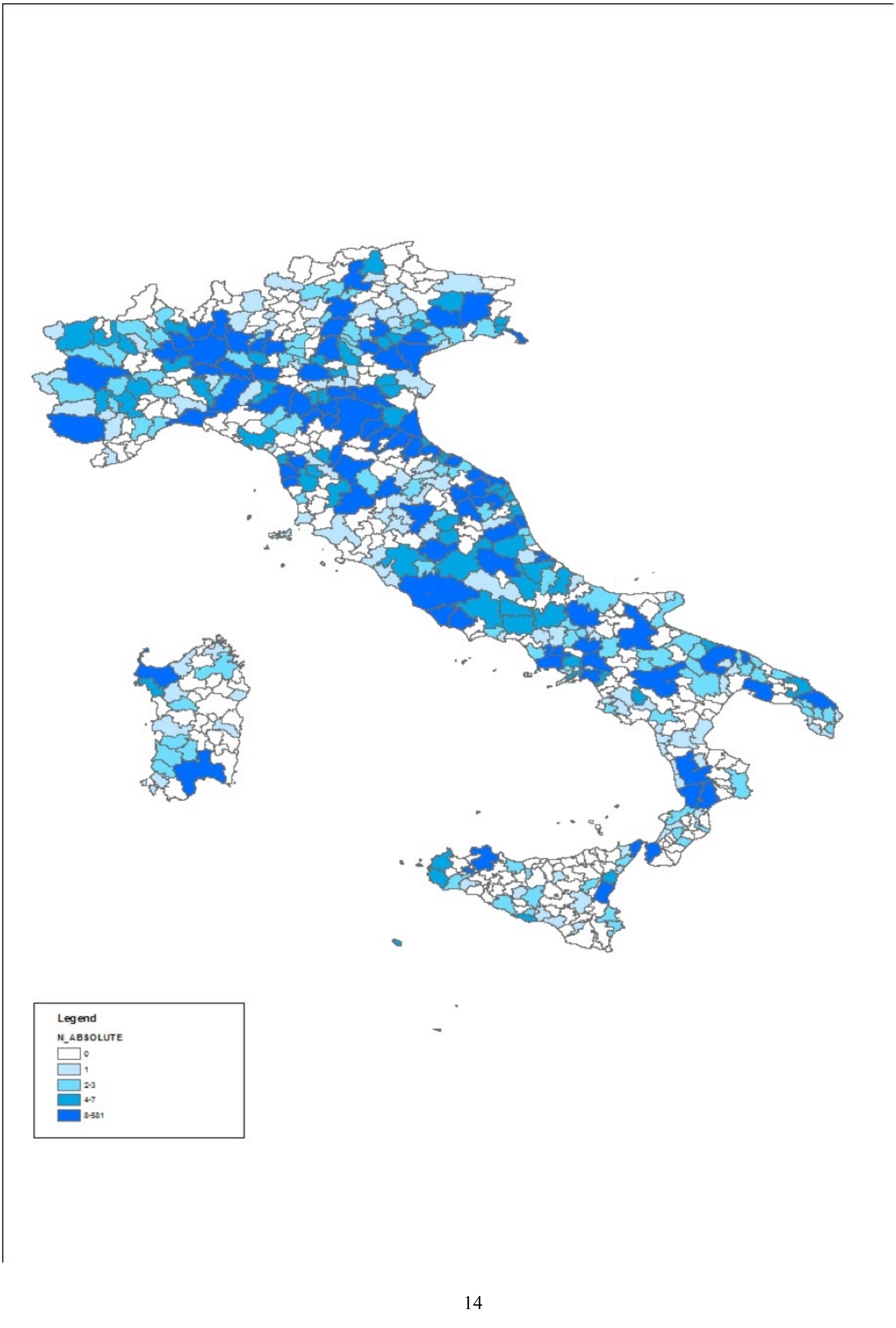
Wieser, R. (2005). Research and development productivity and spillovers: empirical evidence at the firm level. *Journal of Economic Surveys*, 19: 4, 587-621.

Vesper, Karl H., (1990). *New Venture Strategies*. Prentice-Hall, Englewood Cliffs, N.J.

Vivarelli, M. (2004). Are all the Potential Entrepreneurs So Good*?, Small Business Economics*, 23:41-49.

**FIGURES**

**Figure 1. The geographical distribution of innovative start-ups**



**TABLE**

**Table 1. Summary statistics**

| | **MEAN** | **STD. DEV.** | **MIN** | **MAX** |
|---|---|---|---|---|
| N | 6.355 | 30.02 | 0 | 581 |
| # PLANTS | 7,868.1 | 23,532.2 | 214 | 385,491 |
| ID | 0.231 | 0.422 | 0 | 1 |
| UNIV | 0.095 | 0.294 | 0 | 1 |
| HK | 0.084 | 0.023 | 0.031 | 0.176 |
| TRADE | 2.452 | 1.389 | 0 | 4 |
| UNEMP | 0.119 | 0.061 | 0.015 | 0.275 |
| FLOWS | 0.257 | 0.145 | 0.002 | 0.661 |
| SELF_D | 0.810 | 0.066 | 0.625 | 0.945 |
| SELF_S | 0.769 | 0.084 | 0.573 | 0.981 |
| FOR | 6591.2 | 22080.8 | 12 | 370,018 |
| DEN | 205.67 | 292.25 | 10.39 | 3,104.9 |
| VAR | 6.845 | 0.644 | 4.202 | 8.030 |
| RV | 2.223 | 0.333 | 0.985 | 2.980 |
| UV | 4.622 | 0.391 | 2.939 | 5.390 |

**Table 2. Correlation matrix**

| | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] | [11] | [12] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [1] # PLANTS | 1 | | | | | | | | | | | |
| [2] ID | 0.01 | 1 | | | | | | | | | | |
| [3] UNIV | 0.46$^{A}$ | -0.03 | 1 | | | | | | | | | |
| [4] HK | 0.36$^{A}$ | -0.05 | 0.58$^{A}$ | 1 | | | | | | | | |
| [5] TRADE | -0.03 | 0.28$^{A}$ | 0.00 | -0.01 | 1 | | | | | | | |
| [6] UNEMP | -0.05 | -0.32$^{A}$ | -0.04 | -0.10$^{B}$ | -0.25$^{A}$ | 1 | | | | | | |
| [7] FLOWS | 0.21$^{A}$ | 0.27$^{A}$ | 0.19$^{A}$ | 0.15$^{A}$ | 0.26$^{A}$ | -0.48$^{A}$ | 1 | | | | | |
| [8] SELF | 0.19$^{A}$ | -0.28$^{A}$ | 0.28$^{A}$ | 0.26$^{A}$ | -0.21$^{A}$ | 0.12$^{A}$ | -0.08$^{B}$ | 1 | | | | |
| [9] DEN | 0.49$^{A}$ | 0.04 | 0.30$^{A}$ | 0.28$^{A}$ | 0.02 | 0.14$^{A}$ | 0.16$^{A}$ | -0.06 | 1 | | | |
| [10] FOR | 0.13$^{A}$ | 0.45$^{A}$ | 0.09$^{B}$ | 0.15$^{A}$ | 0.36$^{A}$ | -0.70$^{A}$ | 0.40$^{A}$ | -0.24$^{A}$ | 0.05 | 1 | | |
| [11] RV | 0.25$^{A}$ | 0.08$^{B}$ | 0.27$^{A}$ | 0.39$^{A}$ | 0.05 | 0.27$^{A}$ | 0.15$^{A}$ | 0.01 | 0.37$^{A}$ | 0.01 | 1 | |
| [12] UV | 0.32$^{A}$ | 0.25$^{A}$ | 0.37$^{A}$ | 0.52$^{A}$ | 0.30$^{A}$ | -0.18$^{A}$ | 0.43$^{A}$ | -0.02 | 0.29$^{A}$ | 0.34$^{A}$ | 0.58$^{A}$ | 1 |

Notes: all continuous variables are standardized at zero mean and unit variance. $^{a}$ Significant at 1% level; $^{b}$ significant at 5% level.

**Table 3. Model choice: ZINB Vs HNB**

| MODEL WITH VAR | AIC | BIC |
|---|---|---|
| ZINB | 2150.03 | 2428.18 |
| HNB | 2152.12 | 2430.27 |
| MODEL WITH RV AND UV | AIC | BIC |
| ZINB | 2147.66 | 2434.65 |
| HNB | 2148.13 | 2435.11 |

**Table 4. The local determinants of innovative start-up creation: ZINB estimates**

| METHOD: ZINB | ZERO INFLATION | AME | NB |
|---|---|---|---|
| VAR | 0.999 | | 0.448*** |
| | (0.574) | | (0.123) |
| # PLANTS | -31.90** | -3.086 | 0.163** |
| | (15.58) | | (0.059) |
| ID | -1.244 | | -0.215* |
| | (1.792) | | (0.120) |
| UNIV | 2.131 | | 0.570*** |
| | (1.671) | | (0.136) |
| HK | 0.166 | | 0.375*** |
| | (0.681) | | (0.084) |
| TRADE | 0.201 | | 0.003 |
| | (0.293) | | (0.048) |
| UNEMP | -0.718 | | 0.293* |
| | (0.813) | | (0.178) |
| FLOWS | 0.067 | | 0.324*** |
| | (0.528) | | (0.053) |
| SELF | 1.214** | 0.117 | 0.247*** |
| | (0.514) | | (0.053) |
| DEN | 0.586 | | 0.162** |
| | (0.522) | | (0.066) |
| FOR | 0.018 | | 0.362*** |
| | (0.513) | | (0.111) |
| REGIONAL DUMMIES | YES | | YES |
| N OBS. | 611 | | 611 |
| NONZERO OBS. | 350 | | 350 |
| *A* (OVER-DISPERSION) | | 0.213*** | |
| VUONG TEST (ZINB VS NB) | | 5.08*** | |
| LOG PSEUDOLIELIHOOD | | -1012.01 | |

Notes: all continuous variables are standardized at zero mean and unit standard deviation. All the estin constant term. Standard errors are clustered at LLMA level.
*** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

**Table 5. The local determinants of innovative start-up creation: ZINB estimates**

| METHOD: ZINB | ZERO INFLATION | NB | AME |
|---|---|---|---|
| RV | 0.642 | 0.079 | |
| | (0.432) | (0.096) | |
| UV | 0.534 | 0.432*** | 3.647 |
| | (0.560) | (0.119) | |
| REGIONAL DUMMIES | YES | YES | |
| N OBS. | 611 | | |
| NONZERO OBS. | 350 | | |
| *A* (OVER-DISPERSION) | 0.202*** | | |
| VUONG TEST (ZINB VS NB) | 5.06*** | | |
| LOG PSEUDOLIELIHOOD | -1008.83 | | |

Notes: RV and UV are standardized at zero mean and unit standard deviation. All the estimates include a constant term and all the covariates in Table 5. Standard errors are robust to heteroscedasticity. *** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

**Table 6. Unrelated variety and number of innovative start-ups: Lewbel (2012) IV estimates**

| METHOD: LEWBEL IV | (1) | (2) |
|---|---|---|
| UV | 1.578** | 1.814* |
| | (0.778) | (1.055) |
| REGIONAL DUMMIES | YES | YES |
| N OBS. | 611 | 350 |
| CENTERED $R^2$ | 0.932 | 0.935 |
| BREUSCH-PAGAN TEST (P-VALUE) | 0.000 | 0.000 |
| KP WALD F STATISTICS | 11.74 | 14.51 |
| 5% MAX IV RELATIVE BIAS | 21.42 | 21.42 |
| 10% MAX IV RELATIVE BIAS | 11.33 | 11.33 |
| HANSEN J STATISTICS | 31.64 | 31.58 |
| P-VALUE | 0.336 | 0.338 |

Notes: only UV estimated coefficient is reported. *** Significant at 1% level; ** significant at 5% level; * significant at 10% level.

## Appendix: robustness estimates

**Table A1. The local determinants of innovative start-up creation: HNB estimates**

| METHOD: HNB | LOGIT | AME | TRUNCATED NB | AME |
|---|---|---|---|---|
| VAR | -0.253 | | 0.593*** | 8.572 |
| | (0.195) | | (0.163) | |
| # PLANTS | 13.96*** | 1.936 | 0.177*** | 2.553 |
| | (2.678) | | (0.057) | |
| ID | -0.035 | | -0.221* | -3.198 |
| | (0.357) | | (0.124) | |
| UNIV | -0.342 | | 0.591*** | 8.534 |
| | (0.857) | | (0.128) | |
| HK | 0.491** | 0.068 | 0.345*** | 4.988 |
| | (0.193) | | (0.078) | |
| TRADE | -0.082 | | 0.018 | |
| | (0.085) | | (0.044) | |
| UNEMP | 0.455 | | 0.372** | 5.371 |
| | (0.293) | | (0.170) | |
| FLOWS | 0.337* | 0.047 | 0.306*** | 4.416 |
| | (0.179) | | (0.051) | |
| SELF | -0.277* | -0.038 | 0.216*** | 3.122 |
| | (0.145) | | (0.051) | |
| DEN | -0.109 | | 0.112** | 1.619 |
| | (0.210) | | (0.056) | |
| FOR | 0.436** | 0.060 | 0.343*** | 4.953 |
| | (0.212) | | (0.116) | |
| REGIONAL DUMMIES | YES | | YES | |
| N OBS. | 611 | | 611 | |
| NONZERO OBS. | 350 | | 350 | |
| PSEUDO $R^2$ | 0.386 | | 0.246 | |
| *A* (OVER-DISPERSION) | | 0.215*** | | |
| LOG PSEUDOLIKELIHOOD | | -1013.06 | | |

Notes: all continuous variables are standardized, with zero mean and unit standard deviation. All the estimates include a constant term. Standard errors are robust to heteroscedasticity. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

**Table A2. The local determinants of innovative start-up creation: HNB estimates**

| METHOD: HNB | LOGIT | AME | TRUNCATED NB | AME |
|---|---|---|---|---|
| RV | -0.472** | -0.065 | 0.181* | 2.566 |
| | (0.184) | | (0.101) | |
| UV | 0.143 | | 0.509*** | 7.209 |
| | (0.188) | | (0.152) | |
| REGIONAL DUMMIES | YES | | YES | |
| N OBS. | 611 | | 611 | |
| NONZERO OBS. | 350 | | 350 | |
| PSEUDO $R^2$ | 0.385 | | 0.247 | |
| *A* (OVER-DISPERSION) | | 0.201*** | | |
| LOG PSEUDOLIKELIHOOD | | -1005.74 | | |

Notes: all continuous variables are standardized, with zero mean and unit standard deviation. All the estimates include a constant term. Standard errors are robust to heteroscedasticity. *** significant at 1% level; ** significant at 5% level; * significant at 10% level.