Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento Economia e Statistica Cognetti de Martiis

Working Paper Series

06/16

GREEN STARTUPS AND LOCAL KNOWLEDGE BASES: NEWBORN SUPPLIERS OF ENERGY-RELATED TECHNOLOGIES IN ITALIAN PROVINCES

ALESSANDRA COLOMBELLI and
FRANCESCO QUATRARO

brick Bureau of Research on Innovation, Complexity and Knowledge

UNIVERSITÀ DEGLI STUDI DI TORINO

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Green startups and local knowledge bases:

Newborn suppliers of energy-related technologies in Italian Provinces

Alessandra Colombelli[a,c] and Francesco Quatraro[b,c1]

a) DIGEP, Politecnico of Torino (Italy)
b) Department of Economics and Statistics Cognetti de Martiis, University of Torino (Italy)
c) BRICK, Collegio Carlo Alberto (Italy)

ABSTRACT. There is wide consensus about the importance of green technologies for achieving superior economic and environmental performances. The literature on their determinants has neglected the creation of green start-ups as a channel to bring about green technologies in the market. Drawing upon the knowledge spillovers theory of entrepreneurship, we test the relevance of local knowledge stocks, distinguishing between clean and dirty stocks, for the creation of green start-ups. Moreover, the effects of the technological composition of local stocks is investigated, by focusing on technological variety, both related and unrelated, as well as on coherence. Consistently with recent literature, green start-ups are associated to higher levels of variety, pointing to the relevance of diverse and heterogeneous knowledge sources, but in related and complementary technological fields.

Keywords: Green start-ups; New Firm Formation; Energy-related technologies; Knowledge-Spillovers Theory of Entrepreneurship

JEL Classification Codes: L26, M13, R11, O33

[1] Corresponding author: Department of Economics and Statistics Cognetti de Martiis, University of Torino. Lungo Dora Siena 100A, 10153 Torino, Italy. Email: Francesco.quatraro@unito.it.

1 Introduction

The economic analysis of environmental issues has received increasing attention over the last decades. Within the wide body of literature on the subject, the dynamics of the creation of environmental innovations has recently become a key topic. Green technologies are actually regarded as a means of restoring the competitiveness of advanced countries, which has been harmed by the economic crisis (Gilli et al., 2014; Costantini et al., 2013; Cainelli et al., 2013; Ghisetti and Quatraro, 2013; Mazzanti and Zoboli, 2009). Their emergence is indeed supposed to bring about new jobs and new perspectives for economic growth.

The implementation of empirical analyses of eco-innovations (EIs) impacts and determinants has very often focused on the largely established measure of patents, or on the Community Innovation Surveys (CIS). These two approaches allows to appreciating the generation and the adoption of eco-innovation. However, less attention can be found to the main source of innovation that Schumpeter identified in his seminal 1912 book *The theory of Economic Development*, i.e. the entrepreneur. In this perspective, eco-innovation can be brought about in a specific context through the creation of new startup firms involved in the generation and commercialization of technologies improving the environmental performances of the firms that adopt them.

Such missing link is especially problematic as there is increasing consensus on the key role of startups in the introduction of innovation and new technologies in the market, above all when radical technologies are at stake, and their resulting contribution to economic growth (Aghion and Howitt, 1992; Wennekers and Thurik, 1999; Carree and Thurik, 2003; Audretsch et al., 2006). Moreover new firm formation is a determinant of regional growth, cross-regional differences and regional employment dynamics (Fritsch and Schindele, 2011; Dejardin and Fritsch 2011). Therefore understanding the dynamics of creation of green startups can provide useful information on how to boost local development through the interaction of the positive effects of eco-innovations and entrepreneurial dynamics.

This paper aims at filling this gap by linking the analysis of the determinants of eco-innovations to the wide body of literature that has investigated the relationship of entrepreneurship with economic development at the regional level. In this stream of literature, starting by the observation that entrepreneurial activity is geographically clustered, both theoretical and empirical analyses have tried to identify the characteristics and attributes of the local socio-economic systems that may have an impact on new firm formation (Fritsch, 1997; Reynolds et al., 1994; Carlton, 1983; Bartik, 1985; Audretsch and Fritsch, 1994; Feldman 2001; Lee et al. 2004; Colombelli, 2015; Quatraro and Vivarelli, 2015).

In particular, a more recent strand of literature has pointed to the importance of local knowledge spillovers to the entrepreneurial process. A key reference in this domain is the Knowledge Spillovers Theory of Entrepreneurship (henceforth KSTE), conceptualized by Audretsch (1995) and further developed by Audretsch and Lehmann (2005) and Acs et al. (2009), linking knowledge spillovers to new-firm startup activity.

The contribution of our paper to the extant literature is twofold. On the one hand we extend the KSTE to the analysis of green startups, by disentangling the differential impact of 'clean' and 'dirty' knowledge stock. On the other hand we qualify the argument according to which green technologies benefit from heterogenous knowledge sources, by showing that related variety and coherence play an important role.

Our analysis is focused on the patterns of new firm formation in Italian NUTS 3 regions (i.e. the "provincia" level) by using the data on the creation of innovative startups in energy-related technologies (henceforth ERT) within the framework of the new regulation established through the Law Decree n. 179, 2012 October 18th.

This appears an appropriate context for our analysis for different reasons. First, the Italian economy appears to be stuck in mature industries and significantly late from a technological viewpoint, as compared to other most advanced countries, so that our investigation allow us to test the extent to which the relationship between the creation of innovative startups and technological knowledge is shaped by the regional technology context. Second, the Italian case has recently been the object of increasing attention, due to both the availability of emissions levels data at the regional and sectoral level, and to strong regional heterogeneities in environmental performances attention (e.g. Costantini et al., 2013; Ghisetti and Quatraro, 2013 and 2014; Marin and Mazzanti, 2013; Mazzanti and Zoboli, 2009).

2 Theoretical framework

The literature on the determinants and effects of new firm formation has gained momentum in the last decades (Vivarelli, 2013; Quatraro and Vivarelli, 2015). Out the reasons underling such an interest, the importance of entrepreneurs to the innovation process is undoubtedly the most relevant.

More recently, the academic and policy debate on the determinants and effects of innovation has begun to focus more and more on the capacity to reconcile economic and environmental performance through the generation, adoption and diffusion of green technologies. These are actually considered as key to restore the competitiveness of advanced countries harmed by the economic crisis. Their emergence is indeed supposed to create new jobs and introduce new perspectives for economic growth (Crespi et al., 2015). These arguments draw upon the so-called Porter hypothesis (Porter and van der Linde, 1995), according to which innovations aimed at improving firms' environmental performances might also have positive effects on their economic performance due to the enhancement of products and processes engendered by adoption of the innovation.

Given the policy relevance of the phenomenon, based on the so-called double-externality problem, the prevalent interest in the analysis of the determinants of environmental innovation has concerned the extent to which environmental regulation may exert an incentive for firms to introduce innovations, for instance allowing to meeting the polluting standards exogenously set up by policymakers. These studies adopt an induced innovation framework, in which stringent policy frameworks engender additional costs for firms, which increase total production costs by changing the relative factor prices. Firms adopt eco-innovations to save these costs, and in so doing they engender an increase in the derived demand for green technologies (Colombelli et al., 2014; Ghisetti and Quatraro, 2013; Brunnermeier and Cohen, 2003; Rennings and Rammer, 2011; Rennings and Rexhäuser, 2011; for critical reviews of empirical studies see del Rio, 2009 and del Rio et al., 2015). A different and yet related approach to the investigation of the endogenous factors leading to the introduction of eco-innovations can be found in the literature on corporate social responsibility (CSR) (Orlitzky et al. 2011; Hart, 1997).

The extant works on the determinants of green technologies therefore stress on the one hand the effects of environmental regulation, and their impact on firms' economic and financial performances on the other hand. A first systemic overview on this positive relationship is provided by Ambec and Lanoie paper on whether it pay to be green or not (Ambec and Lanoie, 2008). Ghisetti and Rennings (2014) have recently extended the framework of analysis by showing the importance to distinguish between different kinds of eco-innovations when studying their determinants and effects.

The empirical literature has mainly focused on green technologies by using either the Community Innovation Survey (CIS) or patent applications, which are featured as 'green' according to international classification schemes, mainly the WIPO Green Inventory, the OECD EnvTech and the ECLA Y02 class. Less attention is devoted instead to the role of entrepreneurship as a driver of innovations in the realm of environment (Meyskens and Carsrud, 2013; Cohen and Winn, 2007).

Actually entrepreneurs are considered as main agents of change, and in this respect the establishment of new ventures is clearly an important channel through which new green technologies are brought about in the market. Enquiring into the mechanisms of creation of new green startups represents therefore an additional, although less explored, avenue to understand the determinants and effects of green technologies. The grafting of the analysis of the generation eco-innovation onto the so-called knowledge spillovers theory of entrepreneurship (KSTE) would be far reaching in that it would allow to identifying how the formation of green startups is tied to the features of local contexts, both in terms of availability of local stock of knowledge and in terms of scope and complementarity of technological competences accumulated over time (Colombelli, 2015; Colombelli and Quatraro, 2013).

According to the KSTE new knowledge and ideas are one main source of entrepreneurial opportunities (Acs and Armington, 2006; Audretsch et al., 2006). More precisely, new knowledge and ideas created in an incumbent organization, like a firm or a university research laboratory, but left un-commercialized may serve as a source of entrepreneurial opportunities. In this view, the start-up of a new firm is a mechanism for knowledge spillovers from the incumbent organization creating opportunities to the new firm exploiting that opportunities. The KSTE thus suggests that the startup of a new firm is an endogenous response to opportunities generated but not fully exploited by incumbent organizations.

An important implication of the KSTE is that contexts characterized by greater amounts of knowledge generate more entrepreneurial opportunities. Indeed, “contexts rich in knowledge should generate more entrepreneurship, reflecting more extensive entrepreneurial opportunities. By contrast, contexts impoverished in knowledge should generate less entrepreneurship, reflecting less extensive entrepreneurial opportunities” (Audretsch and Keilbach 2007, p. 1249).

In this direction, the investigation of the relationship between the availability of knowledge spillovers and the creation of green start-ups can benefit from insights of the recent literature enquiring into the differential spillover effects from clean and dirty technologies (Dechezelpetre et al., 2013), according to which clean technologies are likely to yield larger spillovers effects than the ‘dirty’ ones.

In view of the previous arguments, we can advance the following hypotheses:

H1. *The amount of knowledge locally available has a positive effect on the creation of 'green' innovative start-ups in a focal province*;

H1a. *Spillovers from clean technologies are stronger than spillovers from 'dirty' ones.*

A large body of empirical analyses has investigated and provided support to the KSTE. In these seminal works, the locally available stock of knowledge actually confirms to be a key variable explaining new firm formation. The local knowledge stock is therein usually proxied by R&D investments (Audretsch and Keilbach 2007; Acs et al. 2009) or by the research efforts carried out in the co-localized universities and research centres (Audretsch and Lehmann 2005; Cassia, Colombelli, Paleari 2009; Cassia and Colombelli 2008; Bonaccorsi et al. 2013; Bonaccorsi et al. 2014). However, these former studies neglect that not only the size of the knowledge stock matters, but also its nature does. The recent evolutionary approaches to economic geography stress indeed the relevance of industrial and technological variety, as well the distinction between related and unrelated variety, in shaping the emergence of new sectors and technologies, also when driven by the creation of new ventures (Boschma, 2005 and 2011; Boschma and Wenting, 2007; Quatraro, 2010; Colombelli et al., 2014). In this line of thought, more recent empirical analyses focus on the effects of knowledge variety on new firm formation (Bae and Koo 2008; Bishop 2012; Colombelli and Quatraro 2013, Colombelli, 2016). These works can be framed into the literature that emphasizes that knowledge spillovers frequently occur across sectors (Jacobs' externalities). In this view, diversity in the local knowledge stock may have a positive impact on the generation of opportunities that entrepreneurs can exploit.

The appreciation of the heterogeneous nature of local knowledge bases makes the investigation of the relationship between knowledge spillovers and green start-ups closely related to the recent stream of literature on the knowledge sources underlying green technologies (Rennings and Rammer, 2009; Horbach et al., 2013; Ghisetti et al., 2015). In this respect Florida (1996), Oltra and Saint Jean (2005a and b), Rennings and Rammer (2009) and De Marchi (2012) provide some insightful evidence about the importance for the emergence of green technologies of the access to different and heterogeneous knowledge sources. Ghisetti et al. (2015) provide robust econometric evidence of the so-called open eco-innovation model (OEIM), according to which the breadth and depth of knowledge sources have a positive effect on the generation of eco-innovation.

In view of these arguments developed above, we can put forth the following hypothesis:

H2. *Knowledge Variety amongst the technological domains featuring the local knowledge base has a positive effect on the creation of new green innovative start-ups in a focal province*

While the extant literature emphasizes the relevance of the diversity of knowledge sources for the generation of eco-innovations, it has paid less attention to the attributes of such heterogeneity. Ghisetti et al. (2015) suggest that green technologies are less path-dependent than other innovations, and therefore their implementation within incumbent firms may involve the ability to deal with skills and competences that are fare away from their existing knowledge base. However, there is no evidence about how the evolutionary patterns of local knowledge bases affect the creation of green startups, i.e. new firms with a green knowledge base. For example, the distinction between related and unrelated variety can prove useful in this context (Frenken et al., 2007; Quatraro, 2010). The very understanding of the features of the local knowledge infrastructure triggering the creation of green innovative startups has to rely on a framework enabling to assess the extent to which they are more likely to emerge in contexts featured by search processes spanning over a wide array of technologies that are loosely or tightly related to one another. The complementarity amongst the technologies observed in a specific area has actually proved to be positively associated to innovation and entrepreneurship at the local level.

Hypothesis 3. *The attributes of the diversity of local knowledge bases matter in shaping the creation of green startups. Related variety and the complementarity amongst knowledge sources are expected to yield a positive impact.*

3 The Italian Context

At the end of 2012 the Italian Ministry of Economic Development approved a Law Decree on “Further urgent measures for Italy’s economic growth”, providing for specific measures which are aimed at promoting the creation and development of innovative start-ups. This was the first time the Italian legislation took this kind of companies into consideration. The law recognizes that start-ups are important for the promotion of sustainable growth, technological development and employment, in particular youth employment, and aims at developing an environment that foster the creation of entrepreneurial opportunities, innovation and social mobility; strengthen the links between universities and businesses; attract to Italy investments and talented people from abroad. Under this law, at the end of 2014 more than 2000 innovative start-ups registered at the Chambers of Commerce in Italy.

In order to be included in the register of “innovative start-ups” and to benefit from governmental incentives, a new company needs to fulfill some requirements. In particular, according to the Law Decree definition, a start-up is a corporation, not listed and subject to the Italian tax law, that has a turnover lower than 5 million euros, is operational for less than 48 months, is owned directly, for at least the 51% by physical subjects, and, most importantly, its social aim is the development of innovative products or services, with an high technological content.

In order to satisfy this latter requirement and to be defined as innovative, the start-up needs to fulfill at least one out of three criteria: either 15% of its costs are related to R&D activities; at least one third of the team is made up of high-qualified members. Finally, the enterprise is the holder, depositary or licensee of a registered patent or the owner of a program for original registered computers.

All the companies included in the register of "innovative start-ups" benefit from the support measures provided by the Law Decree like, for example, the possibility to use start-up's specific flexible employment contracts, to remunerate their team members and the providers of external services with stock options and work for equity, respectively, and to access to incentives for the employment of highly qualified personnel. Moreover, the Law Decree introduces a "fail fast" procedure with the aim to give the entrepreneur the chance to start a new business project as soon as possible.

In addition to the above, the Italian Government in the attempt to stimulate entrepreneurial activities provides some specific measures and incentives for incubators or accelerators that fulfill specific requirements concerning the start-up's physical structures, management, facilities and its track record and also aims at increasing the resources available for venture capital.

Given the peculiarities of the firms included in the Italian register of "innovative start-ups", this appears an appropriate context to test the impact of knowledge spillovers on entrepreneurial activities in energy-related technologies.

4 Data, Methodology and Variables

4.1 Data

Our sample includes 3712 innovative start-ups registered at the Chambers of Commerce in Italy. In particular, we restricted our analysis to companies included in the "innovative start-ups" online directory that registered at the Italian Chamber of Commerce between 2009 and 2015 in 103 Italian NUTS3 regions.

As knowledge spillovers are geographically bounded, we need to focus on a sufficiently narrow definition of region. The unit of analysis in this study is thus the NUTS 3 geographical area. The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU. According to this nomenclature, EU countries are divided into geographical units at three levels of aggregation: NUTS 1, major socio-economic regions; NUTS 2, basic regions for the application of regional policies; NUTS 3, small regions for specific diagnoses. In Italy NUTS 3 regions correspond to administrative units (province) that groups together different neighboring municipalities. Commonly, this administrative unit includes a city and its satellite municipalities. The NUTS 3 geographical area is characterized by the presence of frequent economic interactions. For example, almost each Italian NUTS 3 region hosts a Chamber of Commerce and an employer association. For this reason, this units of analysis is the most appropriate to define the regional boundary of entrepreneurial activities.

In order to analyze the impact of the structure of local knowledge bases on the formation of new firms, we matched data on innovative start-ups, aggregated at the NUTS3 level of analysis, with information contained in the OECD RegPat Database (February 2015) and also with data provided by the Italian institute of statistics (ISTAT), specifically the "Indicatori territoriali per le politiche di sviluppo" (local indicators for development policy).

The OECD RegPat is derived from the Patstat database, which ensures worldwide coverage. These data combine both applications to the EPO and the application to the national patent offices, allowing for going back to 1920 for some patent authorities. This allows for overcoming the traditional limitation of EPO based longitudinal analysis due to its relatively young age.

Patent applications are regionalized at the NUTS 3 level based on inventors' addresses. Applications with more than one inventor residing in different regions have been assigned to each of the regions according to the respective share. Our study is limited to the applications submitted by inventors residing in Italian regions, and uses International Patent Classification (IPC) maintained by the EPO to assign applications to technological classes.

Patents were then defined as being *environmental* on the basis of the World Intellectual Property Organization "WIPO IPC green inventory", an International Patent Classification that identifies patents related to the so-called "Environmentally Sound Technologies" and scatters them into their technology fields, with the *caveat* that it is not the only possible classification of green technologies and, as with other available classifications, it presents some drawbacks (Costantini et al., 2013b)[2]. In particular, consistent with our focus on the determinants of startups in ERT, we

[2] Although interesting, it is out of the scope of the current work to systematically test for the differences that may arise from the choice of classification. We selected the WIPO IPC green inventory since it is currently a wide and well established classification of green technologies. The OECD has indeed also developed the OECD Indicator of Environmental Technologies (OECD, 2011), based on the International Patent Classification (IPC), which features seven environmental areas, i.e. (a) general environmental management, (b) energy generation from renewable and non-fossil sources, (c) combustion technologies with mitigation potential, (d) technologies specific to climate change mitigation, (e) technologies with potential or indirect contribution to emission mitigation, (f) emission abatement and fuel efficiency in transportation, and (g) energy efficiency in buildings and lighting. At the same time, the European

focus on two subgroups of the WIPO Green Inventory, i.e. Energy Conservation and Alternative Energy Production (see the appendix for the correspondence with IPC technological classes).

4.2 Variables

4.2.1 The dependent variable

In order to implement our empirical analysis we took the cumulative sum of innovative startups registering for value added tax (VAT) by NUTS3 region in energy-related technologies (ERT)[3]. These data are provided by the Union of the Chambers of Commerce (Unioncamere) through the Movimprese dataset. A company is in the ERT group if it exclusively develops and commercializes high-value innovative goods and services in the energy field.

It is worth noting that the extant literature proposes two alternative approaches to the measurement of new firm formation, i.e. the ecological and the labor market approach. The ecological approach standardizes figures about new firm creation by using the stock of existing firms, while the labor market approach uses the employment level. These are found to yield very different results when implemented in empirical settings characterized by the same exogenous variables (Audretsch and Fritsch, 1994).

More recently, Audretsch and Lehman (2005) and Bonaccorsi et al. (2013) have assumed that new firms in local contexts could be interpreted as count data. We follow this approach and use the yearly count of the new ERT start-ups in each province ($ERT_{i,t}$) as the dependent variable[4].

Figure 1 shows the distribution of the overall and ERT innovative start-ups across Italian NUTS 3 regions[5].

>>> INSERT FIGURE 1 ABOUT HERE <<<

Patent Office (EPO) is working on completing its own system of classification (ECLA) to assign each patent a green tag, depending on the environmental aim of each patent. So far, EPO allows tagging technologies for adaptation or mitigation to climate change (Y02), in terms of buildings (Y02B), energy (Y02E), transportation (Y02T) and capture, storage sequestration or disposal of GHG (Y02C). More recently, Costantini et al. (2013b) have pointed to the shortcomings of classification methods based on efforts to collect IPCs potentially related to green technologies in one place. Focusing on the biofuels sector, they show that the WIPO Green Inventory is likely to overestimate the number of patents to be assigned due to the fact that IPCs are not specifically designed to identify this narrow and very specific domain. Clinical analysis based on keyword search and validations from experts are likely to yield finer grained classifications. Nonetheless, due to the wide scope of our analysis which encompasses many kinds of green technologies, we will rely on the WIPO Green Inventory.

[3] Data are publicly available at the address http://startup.registroimprese.it/. The data used in this paper are updated to May 2015.

[4] However, we do not deny that local markets are not homogenous with respect to size, and that this can introduce some biases in our results. For this reason, as we specify below, we introduce the employment level in the province among the control variables.

[5] Four provinces in the Sardinia region are not shown in the map, as there are no available data for them. This exclusion however does not alter the results of our empirical analyses.

4.2.2 Key explanatory variables

1.1.1.1 Knowledge Stock

The test of the KSTE involves the use of a measure of local knowledge stock. One can uses either an input or an output measure in this respect. The former would refer to local expenditure for research and development (R&D) as a proxy of the available pool of technological knowledge (Acs et al., 2009). Unfortunately, there are no available data concerning R&D expenditure at the NUTS 3 level in Italy. For this reason we adopt an output measure, i.e. the local knowledge stock (KSTOCK), which is calculated by using patent applications and applying the permanent inventory method as it follows. We calculated the cumulated stock of past patent applications using a rate of obsolescence of 15% per annum:

$$KSTOCK_{i,t} = \dot{h}_{i,t} + (1-\delta)KSTOCK_{i,t-1}, \qquad (1)$$

where $\dot{h}_{i,t}$ is the flow of patent applications and δ is the rate of obsolescence[6], where once again i is the region and t is the time period.

As anticipated in Section 4.1, to test H1a we build the stock of clean knowledge for each region by using the WIPO IPC Green Inventory, by focusing on Energy Conservation and Alternative Energy Production technological fields. We labelled this variable **GT_KSTOCK**. Then we calculated the complement variable **NOGT_KSTOCK** as the stock of patents that are not associated to the those two technological fields.

For what concerns the measurement of the characteristics of the local knowledge base, while some previous studies have used sectoral data (Bishop, 2012) or patent citations (Bae and Koo, 2005), in this paper we build upon the empirical approach put forth in Colombelli (2016) and Colombelli and Quatraro (2013). We use the information contained in patent documents[7] to calculate a number of variables that characterize the local knowledge base based on the complementarity, variety and similarity degree amongst its components. The implementation of knowledge indicators rests on the recombinant knowledge approach.

---

[6] A similar approach is used by Soete et Patel (1985).

[7] The limits of patent statistics as indicators of technological activities are well known. The main drawbacks can be summarized in their sector-specificity, the existence of non-patentable innovations and the fact that they are not the only protecting tool. Moreover the propensity to patent tends to vary over time as a function of the cost of patenting, and it is more likely to feature large firms (Pavitt, 1985; Griliches, 1990). Nevertheless, previous studies highlighted the usefulness of patents as measures of production of new knowledge. Such studies show that patents represent very reliable proxies for knowledge and innovation, as compared to analyses drawing upon surveys directly investigating the dynamics of process and product innovation (Acs et al., 2002). Besides the debate about patents as an output rather than an input of innovation activities, empirical analyses showed that patents and R&D are dominated by a contemporaneous relationship, providing further support to the use of patents as a good proxy of technological activities (Hall et al., 1986).

We consider patents as a proxy for knowledge, and then look at technological classes to which patents are assigned as the constituting elements of its structure. Each technological class *j* is linked to another class *m* when the same patent is assigned to both of them[8]. The higher is the number of patents jointly assigned to classes *j* and *m*, the stronger is this link. Since technological classes attributed to patents are reported in the patent document, we will refer to the link between *j* and *m* as the co-occurrence of both of them within the same patent document[9].

On this basis we calculated the following two key characteristics of regions' knowledge:

a) Knowledge variety (KV) measures the degree of technological diversification of the knowledge base. It is based on the informational entropy index, and is further decomposed in related and unrelated technological variety (RKV and UKV respectively).
b) Knowledge coherence (COH) measures the average degree of complementarity among technologies making up the local knowledge base.

1.1.1.2 Knowledge variety

The measurement of knowledge is based on the information entropy index. Entropy measures the degree of disorder of the system; systems characterized by high entropy are characterized by high degrees of uncertainty (Saviotti, 1988). Informational entropy is a diversity measure showing some interesting properties (Frenken and Nuvolari, 2004) including multidimensionality. This is particularly relevant to our purposes, as it allows us to build up the entropy index on the distribution of the co-occurrences of technological classes in patents, instead of the distribution of the single technological class.

Consider a pair of events (X_l, Y_j), and the probability of their co-occurrence p_{lj}. A two dimensional total variety (*TV*) measure can be expressed as follows:

$$\mathrm{KV} \equiv \mathrm{H}(\mathrm{X},\mathrm{Y}) = \sum_{l}\sum_{j} \mathrm{p}_{lj} \log_2\left(\frac{1}{\mathrm{p}_{lj}}\right) \tag{2}$$

The measure of multidimensional entropy, therefore, draws upon the variety of co-occurrences of technological classes within patent applications, and provides an index of how much the creation of new knowledge is focused on narrower set of possible combinations.

---

[8] In the calculations 4-digits technological classes have been used.
[9] It must be stressed that to compensate for intrinsic volatility of patenting behaviour, each patent application is made last five years in order to reduce the noise induced by changes in technological strategy.

The total index can be decomposed into 'within' and 'between' parts whenever the events being investigated can be aggregated into a smaller number of subsets. Within-group entropy measures the average degree of variety within the subsets; between-group entropy focuses on the subsets, measuring the variety across them. Let the technologies *i* and *j* belong to the subsets *g* and *z* of the classification scheme respectively. If one allows $l \in S_g$ and $j \in S_z$ ($g = 1,\ldots,G$; $z = 1,\ldots, Z$), we can write:

$$P_{gz} = \sum_{l \in S_g} \sum_{j \in S_Z} p_{lj} \tag{3}$$

Which is the probability to observe the couple lj in the subsets *g* and *z*, while the intra subsets variety can be measured as follows:

$$H_{gz} = \sum_{l \in S_g} \sum_{j \in S_z} \frac{p_{lj}}{P_{gz}} \log_2 \left(\frac{1}{p_{lj}/P_{gz}} \right) \tag{4}$$

The (weighted) within-group entropy can be finally written as follows:

$$RKV \equiv \sum_{g=1}^{G} \sum_{z=1}^{Z} P_{gz} H_{gz} \tag{5}$$

Between group (or unrelated variety) can instead be calculated by using the following equation:

$$UKV \equiv H_Q = \sum_{g=1}^{G} \sum_{z=1}^{Z} P_{gz} \log_2 \frac{1}{P_{gz}} \tag{6}$$

Within-group entropy (or related variety) measures the degree of technological differentiation within the macro-field, while between-group variety (or unrelated variety) measures the degree of technological differentiation across macro-fields. The first term on the right-hand-side of equation (7) is the between-entropy, the second term is the (weighted) within-entropy.

We can label between- and within-entropy respectively as *unrelated technological variety* (*UKV*) and *related technological variety* (*RKV*), while total information entropy is referred to as *general technological variety* (Frenken et al., 2007; Boschma and Iammarino, 2009). When variety is high (respectively low), this means that the search process has been extensive (respectively partial).

1.1.1.3 Knowledge coherence

Coherence is defined as the average relatedness or complementarity of a technology chosen randomly within the firm's patent portfolio with respect to any other technology (Nesta and Saviotti, 2006; Nesta, 2008; Quatraro, 2010).

Obtaining the knowledge coherence index requires a number of steps. First of all, we calculate the weighted average relatedness WAR_l of technology l with respect to all other technologies in the regional patent portfolio. This measure builds on the measure of *technological relatedness* among any pair of technologies i and j, τ_{lj} (see Quatraro, 2010).

The weighted average relatedness, WAR_l is then obtained as the degree to which technology l is related to all other technologies $j \in l$ in the region's patent portfolio, weighted by patent count P_{jt}:

$$\mathrm{WAR}_{\mathrm{lt}} = \frac{\sum_{\mathrm{j} \neq \mathrm{l}} \tau_{\mathrm{lj}} \mathrm{P}_{\mathrm{jt}}}{\sum_{\mathrm{j} \neq \mathrm{l}} \mathrm{P}_{\mathrm{jt}}} \qquad (7)$$

The coherence of the region's knowledge base at time t is defined as the weighted average of the WAR_{lt} measure:

$$COH_{\mathrm{t}} = \sum_{\mathrm{l}} \mathrm{WAR}_{\mathrm{lt}} \times \frac{\mathrm{P}_{\mathrm{lt}}}{\sum_{\mathrm{l}} \mathrm{P}_{\mathrm{lt}}} \qquad (8)$$

Note that this index implemented by analysing the co-occurrence of technological classes within patent applications, measures the degree to which the services rendered by the co-occurring technologies are complementary, and is based on how frequently technological classes are combined in use.

4.2.3 Control variables

Besides the effects of the knowledge indicators, we also control for a number of factors that according to the extant literature are likely to affect new firm formation.

First, the possibility to reap the economic benefits stemming from the presence of potentially high demand levels can influence the choice to run a new firm in a specific place. For this reason we control for the effects of agglomeration economies (POP_DENS), proxied by population density, at the NUTS 3 level by dividing the total population at time t in region i by the land use area:

$$POP\_DENS_{i,t} = \frac{POP_{i,t}}{AREA_i}$$

Second, agglomeration economies can also stem from the presence of other firms in the same place, which ensures to some extent the availability of local markets for intermediate goods. In this direction, we also added as a control variable the firm density (FIRM_DENS), calculated as the ratio between the number of registered firms at time *t* in region *i* and the land use area:

$$FIRM\_DENS_{i,t} = \frac{FIRMS_{i,t}}{AREA_i}$$

A complementary measure of prospective economic benefits is also represented by the distance (DIST) of each province *i* from the administrative chief town of the NUTS 2 region (Baptista and Mendonça, 2002; Bonaccorsi et al. 2013).

Third, the creation of new firms can be the outcome of an 'escape from unemployment' strategy. Consistently, we also control for the unemployment rate at the NUTS 3 level (UNEM), calculated as the ration between the count of unemployed people and the count of individuals in the labour force at time *t* in region *i*.

Fourth, we calculated the numbers of incubators (INC) in each province. Actually, business incubators represent a key resource to the creation of new firms, which provide the conditions for successful undertakings and increase the survival likelihood (Colombo and Delmastro, 2002; Auricchio et al., 2014).

Fifth, consistently with the labour market approach to the measurement of entrepreneurship, we include the employment level in the manufacturing sector (MANEMPL) at time *t* in region *i*.

Sixth, a large body of literature has stressed the importance of international trade, and in particular of exports, for the creation of new ventures. Actually, high degrees of internationalization may engender dynamics of 'learning by exporting', based on knowledge about new market and technological opportunities flowing from foreign countries (Blalock and Gertler; 2004; Branstetter 2006; Hessels and van Stel, 2011). For this reason we include in the analysis a variable controlling for the internationalization degree of the NUTS3 region *i* at time *t*. The variable (OPENNESS) is provided by the Italian Institute of Statistics (ISTAT), and is calculated as the share of the value of regional exports in 'dynamic' sectors on total exports[10].

[10] The following Nace Rev. 2 sectors are classified by the ISTAT as 'dynamic': CE-Chemicals; CF-Pharmaceuticals; CI-Computers and electronic and optical products; CJ-Electric apparatus; CL-Transport; M – Professional, scientific and technical activities; R – Arts, entertainment, recreation; S – Other service activities.

Seventh, limited access to financial resources may hamper the entrepreneurial process (Blumberg and Latterie, 2007). Credit rationing is based on information asymmetries, according to which banks may have difficulties in screening investments projects in new ventures, and hence in determining whether a project is a good or bad risk. This engenders a supply shortage for prospective entrepreneurs that can't rely on personal wealth (Stiglitz and Weiss, 1981; Evans and Jovanovic, 1989; Johansson, 2000). In line with this literature, we have included in the econometric model a variable (FIN_SYSTEM) controlling for the quality of the financial markets in NUTS 3 regions, which is proxied by the rate of decay for investments.

Finally, we also control for the quality of the local human by including the share of graduates in the labour force (GRADUATES) and for the presence of high tech business sectors in the local environment (HTKIS).

4.3 Methodology

The basic hypothesis spelt out in section 2 is that the properties of local knowledge base exert an influence on the dynamics creation of ERT start-ups in view of the KSTE. The test of such hypothesis needs for modelling the dependent variable $ERT_{i,t}$ as a function of the characteristics of the knowledge base. The discrete nature and non-negative nature of such dependent variable suggests the adoption of estimation techniques for *count data* models. Out of these models, the equality between conditional variance and conditional mean in the distribution of the dependent variable was violated, suggesting the need for a Negative Binomial class of models instead of a Poisson.

The analysis of the determinants of $ERT_{i,t}$ in our case poses an additional problem which is due to the excess time-region combination for which we observe $ERT_{i,t}=0$. This leads to a situation in which we observe an "excess of zeros" in the dependent variable, and investigation is needed to establish whether the observed zeros are due to the overall absence of innovative startups or to a specific lack of green startups in time-region nonetheless featuring some degree of innovative startups dynamics. For this specificity, we find the zero-inflated negative binomial (ZINB) model is more appropriate to fit our data since it allows empirical frameworks to be modelled in which the excess of zeros in the dependent variable is generated by a different process than count values. This model simultaneously runs two equations: a binary logistical equation to model the zeros in the dependent variable and a proper count data estimation (negative binomial or Poisson) to model the count data dependent variable. In our specification, the LOGIT equation allows us to discriminate between the zeros due to Regions in which some startups are created, but no green startups, and those due to Regions that are not creating any kind of innovative startups, *green* or otherwise. In other words, we based our inflation equation (LOGIT part of the model) on a variable (*TotStartups*) that captures the count of the overall innovative startups (irrespective of whether these were ERT or not) in each time-region combination. The Voung test confirmed the appropriateness of our choice, as reported in the estimation results tables.

To test our hypothesis, the following basic models are specified:

$$ERT_{i,t} = \exp\,(a + \beta_1 KSTOCK_{i,t-n} + \boldsymbol{Z}\gamma + \sum \rho_i + \sum \psi t + \varepsilon_{i,t}) \tag{9a}$$

$$ERT_{i,t} = \exp\,(a + \beta_1 GTKSTOCK_{i,t-n} + \beta_2 NOGTKSTOCK_{i,t-n} + \boldsymbol{Z}\gamma + \sum \rho_i + \sum \psi t + \varepsilon_{i,t}) \tag{9b}$$

$$ERT_{i,t} = \exp\,(a + \beta_1 KV_{i,t-n} + \boldsymbol{Z}\gamma + \sum \rho_i + \sum \psi t + \varepsilon_{i,t}) \tag{10a}$$

$$ERT_{i,t} = \exp\,(a + \beta_{1R} KV_{i,t-n} + \beta_2 UKV_{i,t-n} + \boldsymbol{Z}\gamma + \sum \rho_i + \sum \psi t + \varepsilon_{i,t}) \tag{10b}$$

$$ERT_{i,t} = \exp\,(a + \beta_1 COH_{i,t-n} + \boldsymbol{Z}\gamma + \sum \rho_i + \sum \psi t + \varepsilon_{i,t}) \tag{10c}$$

The error term is decomposed in ρ_i, which is the region fixed effects, the time dummies $\sum\psi t$, and the error component ε_{it}. It must be noted that we run different regressions using different lags for the variables proxying the characteristics of local knowledge base. We report the results of the estimations using five-years lags as they are those that perform better in terms of AIC and BIC. The KSTOCK and the other knowledge-related variables are included separately in the empirical estimations due to the high correlation, as it can be seen in Table 3.

>>> INSERT TABLE 3 ABOUT HERE <<<

The vector $\boldsymbol{Z}$ includes the control variables discussed in Section 4.2.3.. Finally, it is worth noting that all explanatory variables have been transformed by using the inverse hyperbolic sine transformation. In a nutshell, this transformation can be interpreted as a logarithmic transformation, but it is more appropriate when the variables assume value zero for some observations (Burbidge et al. 1988).

5 Econometric results

The results of the econometric estimations of equation (9a) are reported in table 4. The first column shows the fully specified model, while the other columns check for the consistency of the results to the exclusion of key control variables. Let us recall that equation (9a) is intended to test

H1, i.e. the traditional KSTE argument according to which the availability of local knowledge spillovers enhances the creation of innovative startups in regional contexts. The coefficient on KSTOCK is actually positive and significant, providing support to H1. The larger is the knowledge stock available in local contexts, the larger the number of ERT start-ups.

INSERT TABLE 4 ABOUT HERE

This result is important in that it documents once more the relevance of the KSTE approach. However, this would provide little contribution to the extant literature. Nonetheless, in Section 2 we have noticed that the debate about the determinants of green technologies has emphasized the importance of distinguishing between ‘clean’ and ‘dirty’ technologies, suggesting that spillovers from the former are more relevant than those stemming from the former (Dechezlpetre et al., 2013). Table 5 provides therefore a test for H1a, according to which spillovers from clean technologies are more relevant than those generated by the dirty ones. The econometric results are in line with this expectation, as in all of the models reported in the table, the coefficient of GT_KSTOCK is positive and significant, while the one of NOGT_KSTOCK is significant in only one out of four models, although positive. By way of robustness check we have further refined the measurement of the dependent variable by checking, where possible, the activities carried out by each ERT startup in the official list and flagging them as ‘green’ accordingly. The results are reported in the Annex 1. Even if this procedure cannot be fully reliable, due to the firms that do not advertise their activities on the social networks or through a website, the results are in line with the previous estimations.

INSERT TABLE 5 ABOUT HERE

So far, we have provided evidence supporting the hypotheses concerning the importance of the stock of local knowledge, and in particular of the local stock of green knowledge. The following step is to investigate whether the heterogenous nature of local knowledge matters, and to what extent. Actually in Section 2 we have stressed that an increasing body of literature has studied the effect of the breadth and scope of knowledge sources for the introduction of eco-innovations (Ghisetti et al., 2015). We test in this paper the extent to which the technological variety of local knowledge bases that underpin the creation of ERT startups is dispersed across disparate areas of the technological landscape, rather than across loosely related ones.

INSERT TABLE 6 ABOUT HERE

The first column of Table 6 shows the results of the estimation of the impact of knowledge variety (KV) on the generation of ERT startups. Based on the one hand on previous literature about KSTE, and on the other hand on the analyses of knowledge sources for eco-innovations, the empirical results support the hypothesis according to which technological variety positively affects the creation of ERT startups in local contexts. Therefore this result is consistent with H2, according to which ERT innovative start-ups share the basic evidence found about green technologies, i.e. the reliance on knowledge inputs from different and heterogeneous sources (Florida, 1996; Oltra and Saint Jean, 2005; Rennings and Rammer, 2009; Horbach et al., 2013; Montresor et al., 2015).

However, we wish to gain further understanding on the relationship between variety and green startups by investigating whether the kind of implied heterogeneity involves knowledge in related and complementary technological fields or knowledge in apparently disconnected technological fields. The second column of Table 6 shows the differential impact of related and unrelated knowledge variety (RKV and UKV respectively) on ERT[11]. Other things being equal, RKV yields a positive and significant coefficient, while UKV a positive though non significant one. In Column (3) we investigate instead the effect of COH, the coefficient of which turns out to be positive and significant. This would provide support to H3, suggesting that the creation of green startups emerges out of local knowledge bases featured by a high degree of internal coherence, i.e. by the presence of highly complementary technological fields.

6 Conclusions

Innovative start-ups are considered as a powerful instrument for both stagnant economies to recover and developed ones to growth. The financial crisis and the following economic downturn have indeed generated severe resource constraints and unpredictable market conditions that have significantly challenged both developed and emerging countries. Such adverse environmental conditions have fostered a greater need for rethinking the policy agenda both in the EU and overseas to boost economic growth in the years to come. In this vein, at the end of 2012 the Italian Government approved a Law Decree providing specific measures to promote the creation and development of start-ups.

Less attention has been devoted in the empirical literature to the specific case of green start-ups. Actually, being centered around the development and commercialization of eco-innovations, their beneficial impact is related to the win-win framework typical of these technologies. EIs

[11] It is worth recalling that related and unrelated knowledge variety are not opposites, but orthogonal in their meaning (Frenken et al., 2007; Castaldi et al., 2014). In principle, a NUTS 3 region can be characterized by both high RKV and UKV. These would be regions that are diversified into unrelated technological categories while being diversified into many specific classes in each of these categories as well. It is also worth stressing that empirically related and unrelated variety tend to correlate positively (see Table 3; see also Frenken et al., 2007; Quatraro, 2010; Quatraro, 2011; Boschma et al., 2012; Hartog et al., 2012).

actually yield positive effects on both economic and environmental performances. The understanding of their determinants is therefore of paramount importance.

The investigation is based on a theoretical framework combining the KSTE with the specific literature on the determinants of EIs. In particular, we found support for our three hypotheses. The first one states that the availability of local knowledge stock, and ‘clean’ knowledge stock in particular, positively affects the generation of green start-ups. The second one is based on the literature stressing the relevance of diverse knowledge sources to the generation of eco-innovations. Accordingly, we find that technological variety yields positive impact on the generation of green innovative start-ups. Finally, our results show that the kind of technological variety leading to the creation of green startups involves an historical process of knowledge accumulation privileging the combination of related and highly complementary technological fields.

7 References

Acs, Z.J., Braunerhjelm, P., Audretsch, D.B., and Carlsson, B., (2009). The knowledge spillover theory of entrepreneurship. *Small Business Economics*, 32(1), 15–30.

Aghion, P. and Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60, 323-51.

Audretsch, D.B. (1995). *Innovation and Industry Evolution*. Cambridge (Mass), MIT Press.

Audretsch, D.B., Fritsch, M., (1994). The geography of firm births in Germany. *Regional Studies* 28 (4), 359–365.

Audretsch, D.B., Keilbach, M.C. and Lehmann, E.E. (2006). Entrepreneurship and Economic Growth, Oxford, Oxford University Press.

Audretsch, D.B. and Lehmann, E.E. (2005). Does the knowledge spillover theory of entrepreneurship hold for regions? *Research Policy*, 34, pp. 1191–1202

Bartik, T., (1985). Business location decisions in the United States: estimates of the effects of unionization, taxes, and other characteristics of the states. *Journal of Business and Economic Statistics* 3 (January), 16–22.

Cainelli, G., Mazzanti, M. and Zoboli, R. (2013). Environmental performance, manufacturing sectors and firm growth: structural factors and dynamic relationships. *Environmental Economics and Policy Studies* 15, 367-387

Carlton, D.W., (1983). The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables. *Review of Economics and Statistics* 65 (3), 440–449.

Carree, M. and Thurik, A.R. (2006). Understanding the Role of Entrepreneurship for Economic Growth, in Carree, M. and Thurik, A.R. (Eds.), *The Handbook of Entrepreneurship and Economic Growth*. Cheltenham, Elgar, ix-xix

Colombelli A. (2016), The impact of local knowledge bases on the creation of innovative start-ups in Italy, *Small Business Economics*, forthcoming, doi: 10.1007/s11187-016-9722-0.

Costantini, V., Mazzanti, M. and Montini, A. (2013) Environmental performance, innovation and spillovers.Evidence from a regional NAMEA. *Ecological Economics* 89, 101-114.

Crespi, F., Ghisetti, C. and Quatraro, F. (2015). Environmental and innovation policies for the evolution of green technologies: a survey and a test. *Eurasian Business Review* 5, 343-370.

Dejardin M. and Fritsch M. (2011). Entrepreneurial dynamics and regional growth. *Small Business Economics* 36 (4), 377-382.

Feldman, M. (2001). The entrepreneurial event revisited: Firm formation in regional context. *Industrial and Corporate Change*, 10, 861-891.

Fritsch, M. (1997). New firms and regional employment change. *Small Business Economics* 9 (5), 437–448.

Fritsch, M. and Schindele, Y. (2011). The Contribution of New Businesses to Regional Employment—An Empirical Analysis. *Economic Geography*, 87, 153-180.

Ghisetti, C. and Quatraro, F. (2013). Beyond the inducement in climate change: Do environmental performances spur environmental technologies? A regional analysis of cross-sectoral differences. *Ecological Economics* 96, 99-113.

Gilli, M., Mazzanti, M. and Mancinelli, S. (2014). Innovation Complementarity and Environmental Productivity Effects: Reality or Delusion? Evidence from the EU. *Ecological Economics* 103, 56-67.

Lee, S.Y., Florida, R. and Acs, Z., (2004). Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation. *Regional Studies*, 38, 879-891.

Mazzanti, M. and Zoboli, R. (2009) Environmental efficiency and labour productivity: trade-off or joint dynamics?. *Ecological Economics* 68 (4), 1182-1194.

Reynolds, P., Storey, D.J., Westhead, P., (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies* 28 (4), 443–456, July.

Quatraro, F. and Vivarelli, M. (2015). Drivers of entrepreneurship and post-entry performance of newborn firms in developing countries. *World Bank Research Observer* 30, 277-305.

Wennekers, A.R.M. and Thurik, A.R. (1999). Linking entrepreneurship and economic growth. *Small Business Economics*, 13, 27-55.

Figure 1 – Geographical distribution of overall and ERT innovative start-ups

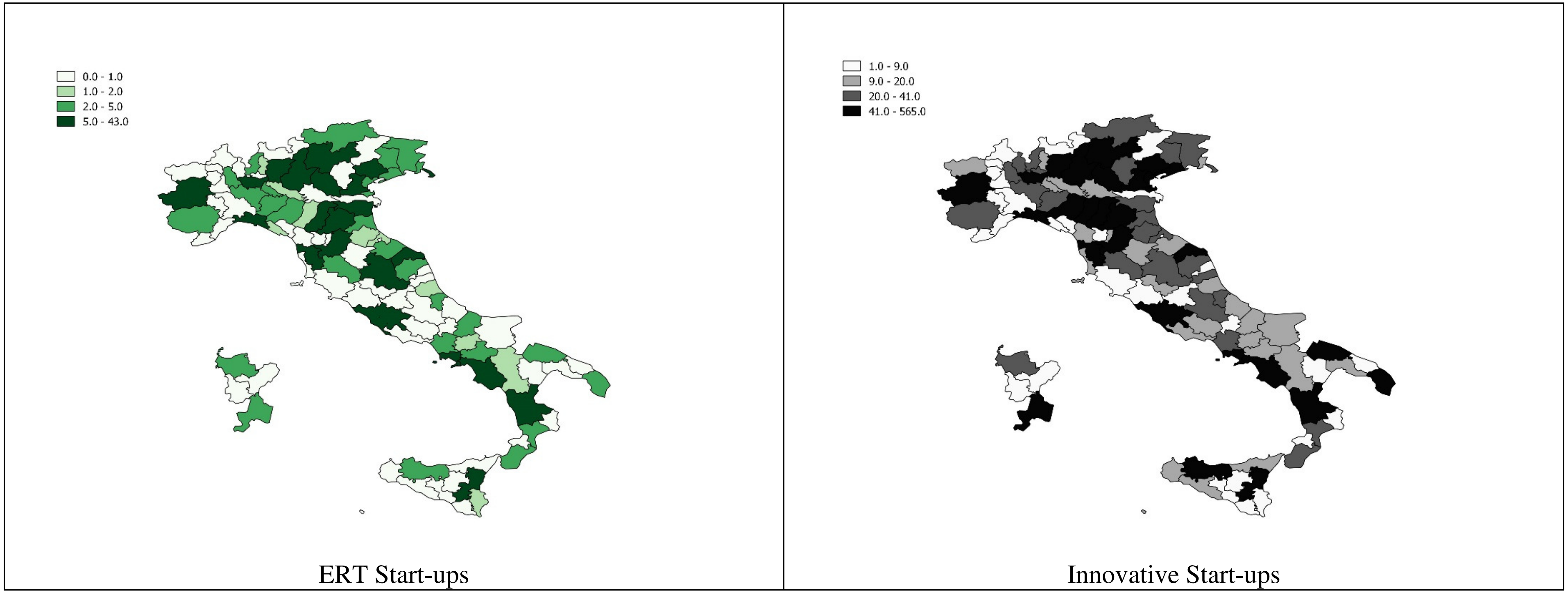

Table 1 – Time Distribution of innovative start-ups

Year	No-ERT	ERT	Total
2009	25	4	29
2010	152	23	175
2011	271	37	308
2012	445	61	506
2013	842	101	943
2014	1,268	165	1,433
2015	473	58	531
Total	**3,476**	**449**	**3,925**

Table 2 - Descriptive statistics

Variable	N	Min	Max	Mean	Sd	Skewness	Kurtosis
ERT	751	0.000	16.000	0.585	1.520	5.429	44.238
GREEN	751	0.000	10.000	0.218	0.685	6.035	64.436
KSTOCK	751	1.095	9.693	5.899	1.451	-0.091	2.725
GT_KSTOCK	746	0.000	4.248	1.093	1.017	0.535	2.440
NOGT_KSTOCK	746	0.000	6.890	3.318	1.300	-0.008	2.538
KV	746	0.000	3.140	2.451	0.474	-2.178	9.902
RKV	746	0.000	2.885	2.055	0.539	-1.710	6.670
UKV	746	0.000	1.778	1.396	0.327	-2.139	8.640
COH	750	-1.642	2.565	-0.634	0.549	1.700	7.387
UNEMP	747	0.019	0.224	0.075	0.040	0.986	3.245
POP_DENS	749	4.580	8.563	5.956	0.756	0.839	4.660
INCUB	751	0.000	2.492	0.396	0.635	1.329	3.567
EXPORT	673	0.652	4.516	3.200	0.708	-0.571	3.422
FINANCE	677	0.160	3.053	1.030	0.389	0.595	4.232
FIRM_DENS	747	0.156	0.418	0.268	0.048	0.407	2.917
GRADUATES	751	0.000	11.273	3.906	4.447	0.312	1.195
HTKIS	483	0.002	0.010	0.005	0.001	0.469	2.988

Table 3 – Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	**ERT**	1.000																
2	**GREEN**	0.737	1.000															
3	**KSTOCK**	0.352	0.284	1.000														
4	**GT_KSTOCK**	0.357	0.263	0.706	1.000													
5	**NOGT_KSTOCK**	0.302	0.213	0.931	0.696	1.000												
6	**COH**	-0.129	-0.137	-0.067	-0.101	-0.029	1.000											
7	**KV**	0.228	0.183	0.831	0.554	0.781	-0.002	1.000										
8	**RKV**	0.246	0.187	0.831	0.548	0.780	-0.013	0.961	1.000									
9	**UKV**	0.154	0.139	0.678	0.471	0.641	0.005	0.807	0.642	1.000								
10	**UNEMP**	-0.028	0.005	-0.549	-0.342	-0.573	-0.077	-0.470	-0.472	-0.425	1.000							
11	**POP_DENS**	0.284	0.213	0.489	0.403	0.475	-0.111	0.410	0.409	0.356	-0.012	1.000						
12	**FIRM_DENS**	-0.121	-0.105	-0.575	-0.388	-0.569	0.059	-0.480	-0.468	-0.477	0.676	-0.231	1.000					
13	**INCUB**	0.376	0.255	0.461	0.389	0.445	-0.075	0.331	0.362	0.204	-0.141	0.296	-0.187	1.000				
14	**GRADUATES**	0.354	0.283	0.332	0.323	0.259	-0.139	0.213	0.235	0.120	0.051	0.149	-0.021	0.419	1.000			
15	**HTKIS**	0.369	0.262	0.581	0.408	0.599	-0.041	0.453	0.470	0.414	-0.492	0.548	-0.619	0.301	0.088	1.000		
16	**EXPORT**	0.104	0.089	0.258	0.106	0.218	-0.139	0.175	0.184	0.145	-0.096	0.070	-0.127	0.287	0.199	0.108	1.000	
17	**FINANCE**	0.014	0.021	-0.336	-0.274	-0.380	-0.115	-0.337	-0.331	-0.286	0.488	-0.158	0.402	-0.161	-0.038	-0.224	-0.050	1.000

Table 4 - Zero-inflated Negative Binomial Estimation, overall Knowledge Stock

	(1)	(2)	(3)	(4)
KSTOCK	0.4437***	0.6178***	0.4331***	0.4349***
	(0.1207)	(0.1099)	(0.1197)	(0.1154)
UNEMP	-4.9289	-5.2959	-5.0528	-5.1224
	(5.0676)	(5.2074)	(5.0514)	(4.9983)
POP_DENS	-0.0352	-0.1118	-0.0244	
	(0.1414)	(0.1428)	(0.1397)	
INCUB	0.0383	0.1130	0.0314	0.0240
	(0.1406)	(0.1453)	(0.1409)	(0.1284)
EXPORT	0.1246	0.1630	0.1319	0.1181
	(0.1397)	(0.1407)	(0.1391)	(0.1372)
FINANCE	-0.2629	-0.2263	-0.2678	-0.2566
	(0.2809)	(0.2914)	(0.2807)	(0.2796)
FIRM_DENS	-1.9575	-2.5226		-1.8657
	(3.0434)	(3.1029)		(3.0168)
GRADUATES	0.0702***		0.0709***	0.0712***
	(0.0233)		(0.0233)	(0.0229)
HTKIS	106.1463	105.9184	112.1900	101.3632
	(80.2250)	(81.5178)	(79.5216)	(77.8890)
Constant	-1.7907	-2.0823	-2.3338**	-1.9179
	(1.3985)	(1.4024)	(1.1151)	(1.3012)
inflate				
TotStartups	-0.3384***	-0.3258***	-0.3431***	-0.3387***
	(0.0980)	(0.0903)	(0.1003)	(0.0982)
Constant	0.9993***	1.0765***	1.0222***	1.0003***
	(0.3748)	(0.3476)	(0.3744)	(0.3750)
lnalpha	-3.6078**	-3.1764**	-3.6533*	-3.6127**
	(1.8247)	(1.3039)	(1.9296)	(1.8398)
N	479	479	479	479
AIC	858.2986	865.4339	856.7147	856.3609
BIC	1004.3082	1007.2717	998.5526	998.1987
Log Lik	-394.1493	-398.7169	-394.3574	-394.1805
McFadden's R2	0.2946	0.2864	0.2942	0.2946
Vuong test	2.7278	2.9752	2.7184	2.7311

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5- Zero-inflated Negative Binomial Estimation, clean vs. 'dirty' knowledge stock

	(1)	(2)	(3)	(4)
GT_KSTOCK	0.2470**	0.2813**	0.2455**	0.2437**
	(0.1154)	(0.1215)	(0.1150)	(0.1147)
NOGT_KSTOCK	0.2139	0.3663***	0.2144	0.2066
	(0.1484)	(0.1420)	(0.1482)	(0.1457)
UNEMP	-6.6151	-7.5750	-6.6368	-6.8149
	(5.1497)	(5.3550)	(5.1433)	(5.0802)
POP_DENS	-0.0368	-0.1175	-0.0333	
	(0.1422)	(0.1459)	(0.1411)	
INCUB	0.0112	0.1361	0.0067	-0.0032
	(0.1422)	(0.1470)	(0.1412)	(0.1309)
EXPORT	0.1929	0.2363	0.1957	0.1849
	(0.1434)	(0.1460)	(0.1429)	(0.1399)
FINANCE	-0.3566	-0.2841	-0.3598	-0.3479
	(0.2887)	(0.3044)	(0.2880)	(0.2864)
FIRM_DENS	-0.7318	-0.8262		-0.6560
	(2.9853)	(3.0790)		(2.9651)
GRADUATES	0.0806***		0.0807***	0.0816***
	(0.0226)		(0.0226)	(0.0223)
HTKIS	101.9781	101.2323	103.7120	97.5570
	(80.1155)	(82.0580)	(79.7037)	(78.2465)
Constant	-0.3038	-0.1452	-0.5238	-0.4604
	(1.3523)	(1.3745)	(1.0105)	(1.2073)
inflate				
TotStartups	-0.3566***	-0.3627***	-0.3576***	-0.3567***
	(0.0956)	(0.1028)	(0.0961)	(0.0959)
Constant	1.1787***	1.3145***	1.1845***	1.1781***
	(0.3554)	(0.3381)	(0.3549)	(0.3558)
lnalpha	-4.5842	-3.3648**	-4.6541	-4.5945
	(4.2889)	(1.5175)	(4.5948)	(4.3478)
N	477	477	477	477
AIC	854.5880	865.6121	852.6483	852.6552
BIC	1004.6186	1011.4752	998.5114	998.5183
Log Lik	-391.2940	-397.8060	-391.3241	-391.3276
McFadden's R2	0.2969	0.2852	0.2968	0.2968
Vuong test	2.8828	3.1861	2.8833	2.8806

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6 - Zero-inflated Negative Binomial Estimation, Knowledge variety and coherence

	(1)	(2)	(3)
KV	1.1925***		
	(0.4419)		
RKV		0.7500**	
		(0.3473)	
UKV		0.5742	
		(0.4883)	
COH			0.3510**
			(0.1755)
UNEMP	-6.0588	-6.0690	-8.2520
	(5.0962)	(5.0694)	(5.2861)
POP_DENS	0.0033	0.0084	0.2137
	(0.1432)	(0.1427)	(0.1423)
INCUB	0.1870	0.1818	0.2048
	(0.1364)	(0.1363)	(0.1391)
EXPORT	0.1207	0.1176	0.1206
	(0.1405)	(0.1400)	(0.1413)
FINANCE	-0.1698	-0.2123	-0.2129
	(0.2820)	(0.2981)	(0.2741)
FIRM_DENS	-1.0159	-0.8018	-1.6723
	(3.0638)	(3.0702)	(3.1080)
GRADUATES	0.0894***	0.0903***	0.1079***
	(0.0225)	(0.0224)	(0.0213)
HTKIS	122.5471	123.7226	98.0689
	(79.8368)	(79.2425)	(82.3373)
Constant	-2.7375*	-2.1686	-0.1181
	(1.6374)	(1.5354)	(1.3485)
inflate			
TotStartups	-0.3430***	-0.3390***	-0.3860***
	(0.0910)	(0.0891)	(0.1000)
Constant	1.0997***	1.1057***	1.3543***
	(0.3598)	(0.3548)	(0.3417)
lnalpha	-3.3639**	-3.5650**	-3.3787**
	(1.4641)	(1.7974)	(1.4939)
N	482	482	489
AIC	865.6473	867.5439	870.8560
BIC	1011.8753	1017.9499	1017.5886
Log Lik	-397.8236	-397.7720	-400.4280
McFadden's R2	0.2896	0.2897	0.2884
Vuong test	3.1808	3.1300	3.5271

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Annex 1 - Zero-inflated Negative Binomial Estimation, subsample of 'green' ERT

	(1)	(2)	(3)	(4)
GT_KSTOCK	0.3456*	0.3388*	0.3328*	0.3535*
	(0.1985)	(0.2027)	(0.1965)	(0.1970)
NOGT_KSTOCK	0.0232	0.2353	0.0360	0.0386
	(0.2515)	(0.2285)	(0.2492)	(0.2469)
UNEMP	-9.2592	-12.0753	-9.6409	-9.1663
	(9.1453)	(9.1145)	(9.0944)	(9.1429)
POP_DENS	0.0868	0.0029	0.0891	
	(0.2504)	(0.2509)	(0.2488)	
INCUB	-0.2248	-0.0683	-0.2138	-0.1888
	(0.2612)	(0.2453)	(0.2640)	(0.2392)
EXPORT	0.3326	0.4167	0.3490	0.3546
	(0.2706)	(0.2693)	(0.2690)	(0.2631)
FINANCE	-0.6352	-0.4890	-0.6281	-0.6540
	(0.4971)	(0.4925)	(0.4993)	(0.4943)
FIRM_DENS	-3.6088	-2.9806		-3.6334
	(5.5538)	(5.7633)		(5.5425)
GRADUATES	0.0944**		0.0928**	0.0922**
	(0.0385)		(0.0384)	(0.0378)
HTKIS	1.2212	20.7966	7.4577	11.8968
	(139.5579)	(138.5253)	(138.9224)	(135.8437)
Constant	-0.0332	-0.3018	-1.1047	0.3170
	(2.4006)	(2.3969)	(1.7350)	(2.1764)
Inflate				
TotStartups	-0.4487**	-0.4752**	-0.4772**	-0.4542**
	(0.1761)	(0.2084)	(0.2058)	(0.1818)
constant	2.0798***	2.2240***	2.1204***	2.0890***
	(0.5983)	(0.6051)	(0.6247)	(0.6011)
lnalpha	-4.1805	-3.1438	-3.5959	-4.0871
	(6.1691)	(2.4137)	(3.5665)	(5.6511)
N	477	477	477	477
AIC	474.6049	479.0743	473.0286	472.7250
BIC	624.6355	624.9374	618.8917	618.5881
Log Lik	-201.3024	-204.5372	-201.5143	-201.3625
McFadden's R2	0.3282	0.3174	0.3275	0.3280
Vuong test	2.5342	2.6165	2.5525	2.5446

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Annex 2 - WIPO Green Inventory, List of Technological Classes

TOPIC	IPC
ALTERNATIVE ENERGY PRODUCTION	
Bio-fuels	
Solid fuels	C10L 5/00, 5/40-5/48
Torrefaction of biomass	C10B 53/02
	C10L 5/40, 9/00
Liquid fuels	C10L 1/00, 1/02, 1/14
Vegetable oils	C10L 1/02, 1/19
Biodiesel	C07C 67/00, 69/00
	C10G
	C10L 1/02, 1/19
	C11C 3/10
	C12P 7/64
Bioethanol	C10L 1/02, 1/182
	C12N 9/24
	C12P 7/06-7/14
Biogas	C02F 3/28, 11/04
	C10L 3/00
	C12M 1/107
	C12P 5/02
From genetically engineered organisms	C12N 1/13, 1/15, 1/21, 5/10, 15/00 A01H
Integrated gasification combined cycle (IGCC)	C10L 3/00
	F02C 3/28
Fuelcells	H01M 4/86-4/98, 8/00-8/24, 12/00-12/08
Electrodes	H01M 4/86-4/98
Inert electrodes with catalytic activity	H01M 4/86-4/98
Non-activeparts	H01M 2/00-2/04 , 8/00-8/24
Within hybridcells	H01M 12/00-12/08
Pyrolysis or gasification of biomass	
	C10B 53/00
	C10J
Harnessing energy from manmade waste	
Agricultural waste	C10L 5/00
Fuel from animal waste and crop residues	C10L 5/42, 5/44
Incinerators for field, garden or wood waste	F23G 7/00, 7/10
Gasification	C10J 3/02, 3/46
	F23B 90/00
	F23G 5/027
Chemicalwaste	B09B 3/00
	F23G 7/00
Industrial waste	C10L 5/48
	F23G 5/00, 7/00
Using top gas in blast furnaces to power pig-iron production	C21B 5/06
Pulp liquors	D21C 11/00
Anaerobic digestion of industrial waste	A62D 3/02
	C02F 11/04, 11/14
Industrial wood waste	F23G 7/00, 7/10
Hospital waste	B09B 3/00
	F23G 5/00
Landfill gas	B09B
Separation of components	B01D 53/02, 53/04, 53/047, 53/14, 53/22, 53/24
Municipal waste	C10L 5/46
	F23G 5/00
Hydroenergy	
Water-power plants	E02B 9/00-9/06
Tide or wave power plants	E02B 9/08
Machines or engines for liquids	F03B
	F03C
Using wave or tide energy	F03B 13/12-13/26
Regulating, controlling or safety means of machines or engines	F03B 15/00-15/22
Propulsion of marine vessels using energy derived from water movement	B63H 19/02, 19/04
Ocean thermal energy conversion (OTEC)	F03G 7/05
Wind energy	F03D
Structural association of electric generator with mechanical driving motor	H02K 7/18
Structural aspects of wind turbines	B63B 35/00
	E04H 12/00
	F03D 11/04
Propulsion of vehicles using wind power	B60K 16/00
Electric propulsion of vehicles using wind power	B60L 8/00
Propulsion of marine vessels by wind-powered motors	B63H 13/00
Solar energy	
Photovoltaics (PV)	
Devices adapted for the conversion of radiation energy into electrical energy	H01L 27/142, 31/00-31/078 H01G 9/20
	H02N 6/00
Using organic materials as the active part	H01L 27/30, 51/42-51/48
Assemblies of a plurality of solar cells	H01L 25/00, 25/03, 25/16, 25/18, 31/042
Silicon; single-crystal growth	C01B 33/02 C23C 14/14, 16/24 C30B 29/06
Regulating to the maximum power available from solar cells	G05F 1/67
Electric lighting devices with, or rechargeable with, solar cells	F21L 4/00
	F21S 9/03

Charging batteries	H02J 7/35

Dye-sensitised solar cells (DSSC)	H01G 9/20
	H01M 14/00
Use of solar heat	F24J 2/00-2/54
For domestic hot water systems	F24D 17/00
For space heating	F24D 3/00, 5/00, 11/00, 19/00
For swimming pools	F24J 2/42
Solar updraft towers	F03D 1/04, 9/00, 11/04 F03G 6/00
For treatment of water, waste water or sludge	C02F 1/14
Gas turbine power plants using solar heat source	F02C 1/05
Hybrid solar thermal-PV systems	H01L 31/058
Propulsion of vehicles using solar power	B60K 16/00
Electric propulsion of vehicles using solar power	B60L 8/00
Producing mechanical power from solar energy	F03G 6/00-6/06
Roof covering aspects of energy collecting devices	E04D 13/00, 13/18
Steam generation using solar heat	F22B 1/00
	F24J 1/00
Refrigeration or heat pump systems using solar energy	F25B 27/00
Use of solar energy for drying materials or objects	F26B 3/00, 3/28
Solar concentrators	F24J 2/06
	G02B 7/183
Solar ponds	F24J 2/04
Geothermal energy	
Use of geothermal heat	F01K
	F24F 5/00
	F24J 3/08
	H02N 10/00
	F25B 30/06
Production of mechanical power from geothermal energy	F03G 4/00-4/06, 7/04
Other production or use of heat, not derived from combustion, e.g. natural heat	F24J 1/00, 3/00, 3/06
Heat pumps in central heating systems using heat accumulated in storage masses	F24D 11/02
Heat pumps in other domestic- or space-heating systems	F24D 15/04
Heat pumps in domestic hot-water supply systems	F24D 17/02
Air or water heaters using heat pumps	F24H 4/00
Heat pumps	F25B 30/00
Using waste heat	
To produce mechanical energy	F01K 27/00
Of combustion engines	F01K 23/06-23/10
	F01N 5/00
	F02G 5/00-5/04
	F25B 27/02
Of steam engine plants	F01K 17/00, 23/04
Of gas-turbine plants	F02C 6/18

As source of energy for refrigeration plants	F25B 27/02
For treatment of water, waste water or sewage	C02F 1/16
Recovery of waste heat in paper production	D21F 5/20
For steam generation by exploitation of the heat content of hot heat carriers	F22B 1/02
Recuperation of heat energy from waste incineration	F23G 5/46
Energy recovery in air conditioning	F24F 12/00
Arrangements for using waste heat from furnaces, kilns, ovens or retorts	F27D 17/00
Regenerative heat-exchange apparatus	F28D 17/00-20/00
Of gasification plants	C10J 3/86
Devices for producing mechanical power from muscle energy	F03G 5/00-5/08

ENERGY CONSERVATION	
Storage of electrical energy	B60K 6/28
	B60W 10/26
	H01M 10/44-10/46
	H01G 9/155
	H02J 3/28, 7/00, 15/00
Power supply circuitry	H02J
With power saving modes	H02J 9/00
Measurement of electricity consumption	B60L 3/00
	G01R
Storage of thermal energy	C09K 5/00
	F24H 7/00
	F28D 20/00, 20/02
Low energy lighting	
Electroluminescent light sources (e.g. LEDs, OLEDs, PLEDs)	F21K 99/00
	F21L 4/02
	H01L 33/00-33/64, 51/50 H05B 33/00
Thermal building insulation, in general	E04B 1/62, 1/74-1/80, 1/88, 1/90
Insulating building elements	E04C 1/40, 1/41, 2/284-2/296
For door or window openings	E06B 3/263
For walls	E04B 2/00
	E04F 13/08
For floors	E04B 5/00
	E04F 15/18
For roofs	E04B 7/00
	E04D 1/28, 3/35, 13/16
For ceilings	E04B 9/00
	E04F 13/08
Recovering mechanical energy	F03G 7/08
Chargeable mechanical accumulators in vehicles	B60K 6/10, 6/30
	B60L 11/16