Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

Social innovation in the US high-tech industries: its core business and main drivers of innovation

Alessandro Marra
Università degli Studi "G. d'Annunzio" Chieti - Pescara
Economics
amarra@luiss.it

Riccardo Maiolini
LUISS Guido Carli
CeRIIS International Center for Social Innovation
rmaiolini@luiss.it

Cristiano Baldassarri
LTKA
-
baldassarri@ltka.org

Abstract

Social Innovation (SI) is a fuzzy concept, difficult to define and circumscribe because of its complexity and novelty. Our purpose in this paper is to give substance to SI by investigating on a large-n base all social and innovative activities carried out by the US high-tech community. Section 1 introduces. Section 2 provides an empirical-based literature review on SI. Section 3 illustrates the methodology and the dataset. Section 4 offers an observational study of SI in US high-tech. Section 5 proposes the network analysis based on metadata (tags on products, technologies and markets) to capture the core business and main drivers of SI. Section 6 concludes.

Social innovation in the US high-tech industries: its core business and main drivers of innovation

Abstract

Social Innovation (SI) is a fuzzy concept, difficult to define and circumscribe because of its complexity and novelty. Our purpose in this paper is to give substance to SI by investigating on a large-n base all social and innovative activities carried out by the US high-tech community. Section 1 introduces. Section 2 provides an empirical-based literature review on SI. Section 3 illustrates the methodology and the dataset. Section 4 offers an observational study of SI in US high-tech. Section 5 proposes the network analysis based on metadata (tags on products, technologies and markets) to capture the core business and main drivers of SI. Section 6 concludes.

1. Introduction

Social Innovation (SI) is a fuzzy concept, difficult to define and circumscribe because of its complexity and novelty. In literature a consensus is emerging on the meaning of SI, but as far as we know a few empirical studies exist, most of which are based on a "case by case" approach (see Bulut et al., 2013a). This paper intends to follow up the call launched by Phillips (2011) and Cajaiba-Santana (2014) for a better definition of the boundaries of SI by providing some empirical evidence on the phenomenon under observation in high tech industries.

According to Mulgan (2006) SI refers to innovative activities and services that are motivated by the achievement of a social need. An innovation is "social" to the extent that it does not necessarily mean good, but is socially desirable (Howaldt and Schwarz, 2010) and as long as satisfies social needs not perceived as relevant by the market (Mulgan et al., 2007). Phills et al. (2008) put emphasis on the term "social", which describes a kind of value that is distinct from financial or economic value, defined as *"the creation of benefits or reductions of costs for society—through efforts to address social needs and problems—in ways that go beyond the private gains and general benefits of market activity"* (Phills et al., 2008: 38). Auerswald (2009) defines SI as *"a novel solution to a social problem that is more effective, efficient, sustainable, or just than existing solutions and for which the value created accrues primarily to society as a whole rather than private individual"* (Auerswald 2009: 52).

According to others, SI implies social interactions (see Marcy and Mumford 2007, Mumford 2002). Hamalainen and Heiscala (2007) describe SI as a way to address needs or resolve difficulties through new ideas and new kind of social structures. This induces a shift towards a demand-pull model where adopted solutions are directed to a large number of individuals, who participate to its development and fruition (Eikins 1992; Guida and Maiolini 2013). Hubert (2010) effectively sums up affirming that SI is social both in its needs and means.

The above theoretical definitions serve as benchmark over which to build (from a practical point of view) our own understanding of SI, developed through a large-n base investigation of all social and innovative activities carried out by the US high-tech community.

The structure of the paper is as follows. Section 2 provides an empirical literature review. Section 3 illustrates the methodology and all qualitative and quantitative information collected from CrunchBase, the world's most comprehensive database of

digital and high-tech start-up activities. Section 4 presents the main figures and trends on SI in the US high tech industries: number of start-ups, employees, total funding raised, industries involved and most recurrent tags. Moreover, a georeferentiation of SI is provided by looking at cities with the largest proportion of social and innovative activities. Section 5 proposes a network analysis based on metadata (tags on products, technologies and markets) to capture the core business and main drivers of SI. Section 6 concludes.

2. An empirical literature review on SI

A few empirical studies exist in literature on SI, most of which are based on a "case by case" approach (Murray et al., 2008; Mumford and Moertl, 2003). Basically, two fields of research has been identified: the former is focused on the identification of factors that facilitate the emersion of social and innovative ideas (Dumford and Moerti, 2003), while the latter looks at the dissemination of innovation, where the "social" dimension is the driver to increase the participation in social issues (Kirwan et al., 2013).

Dumford and Moerti (2003) describe the generation of the ideas underlying SI, the development and refinement of these ideas, and the social factors that promote the implementation of these innovations. Bulut et al. (2013a, 2013b) investigate the emersion of SI from a social perspective: in the former the authors elaborate on the individual propensity of SI creation, through the elaboration of a SI scale, while in the latter examine the causal relationship between SI and technological innovation. It emerges the "relativity" of SI and the fact that it is transversal to different geographical and socio-economic activities.

Kirwan et al. (2013) focus on the decision making process to solve social issues. The authors investigate the relationship between SI and economic and social development. According to them, SI creates benefits in two ways: origin of innovative solutions and new ways to diffuse and replicate best practices. Maruyama et al. (2007) propose a case study to explore "*the rules of risk–benefit distribution and the roles of social actors*" in the elaboration of SI dynamics: citizens played a relevant role in the introduction of photovoltaic systems in Japan and this shows how an individual choice can generate consequences at a societal level and among communities. Rodima-Taylor (2012) study the role of informal associations inside local communities in

Africa, where the diffusion of SI is driven by the informal networks that amalgamate different opinions and drive the identification of shared and innovative solutions.

SI is transversal to many sectors and for this reason is not easy to define. A few studies, elaborated by experts and practitioners, attempt to identify the boundaries of SI. According to the European Commission (EC) and the Bureau of Policy Advisers (BEPA), for example, "*the concept [of SI] currently draws from four theoretical sources: innovation, social investment, change and open society*" (BEPA, 2010: 32). In other words, SI is related to all social and societal demands or societal challenges where any sort of innovation can procure an improvement based on newness and progress. From this perspective, SI takes place in case of innovative changes with deep implications on migration, unemployment, poverty, ageing and health, climate change, education, discrimination, passive dependency and welfare, urban regeneration, social economy, incubation and workplace innovation.

3. Methodology and dataset

3.1. The methodology

Network analysis is a well known and used methodology founded by sociologists and researchers in social psychology and then further developed in collaboration with mathematicians and statisticians to a point where it is now considered to be technically sophisticated and is currently employed in several disciplines, such as economics, marketing and industrial engineering.

Economic tools are generally highly regarded and network analysis has many potential uses within the field of economic analysis –which is one of the reasons it has been so widely employed. The two key areas of application are the understanding of the way networks influence economic activity, such as research and development and patent activity, and its ability to reveal network formation and network influence. Clearly its versatility across different sectors and complementary approach with other disciplines are factors in its wide usage (for an overview of network applications in economics: Bloch, 2004; Jackson, 2006, 2011).

Several studies have investigated R&D and innovation activities by referring to network analysis. Cantner and Graf (2006) applied network analysis to describe the evolution of the innovators network in Jena, Germany between 1995 and 2001. Owen-Smith et al. (2002) looked at R&D cooperation and compared the organization of

4

scientific research in the US and Europe by employing network analysis. Breschi and Lissoni (2003) and Singh (2003) found that social rather than geographical proximity is relevant for knowledge spillovers. Balconi et al. (2004) analysed inventor networks resulting from common team-membership in patenting, focusing on the specific role of academic inventors in different technological classes.

Innovation in high-tech industries is rapidly evolving and to keep up with emerging products, needs and technologies and changing business strategies the proposed network analysis allows for the identification of the pattern of technological change carried out by innovative companies in a given time period.

Metadata (i.e., tags) on products, markets and technologies enables investigation into technological innovation and complementarities in high-tech companies. In the proposed analysis the nodes of the network represent tags and the co-occurrence in one or more companies that use different tags is depicted through the edges of the network. For example tags A and B are linked in the network if these coexist in the same company and the weight is heavier if the number of companies in which the two tags coexist is larger. Therefore for tags A and B, the weight of the edge A-B is 5 since these coexist in five different companies and the weight of the edge A-C is 2 since these coexist in two different companies and so on (Table 1).

Table 1: Row data (example)

	Name	Tags
1	Company$_1$	A, B, C, D
2	…	A, B, E, F
3	…	A, B
4	…	A, B, C, G, H, I
5	Company$_n$	A, B, L, M, N

Edges represent relationships and complementarities between two nodes. In the same company the heavier or more weighted edges represent the actual links between two tags over which innovative start-ups usually invest tangible and intangible resources and typically drive their business. Weaker or fainter edges originating from a specified node also bring information and can be interpreted in terms of potential for horizontal and/or vertical expansion and collaboration between companies that do not exploit the

same technological complementarities. This last view is at the basis of the cluster investigation that is further discussed in Section 3.

The proposed methodology is similar to the first of the two approaches followed by Cantner and Graf (2006) in their investigation on the innovators network in Jena. The approach is called "technological overlap" and consists in linking innovators by their technological knowledge. Therefore, the more fields of research the innovators have in common, the closer they are related. The second approach followed in Cantner and Graf (2006) is related to the notion that knowledge transfer exists through personal relationships as firms and research institutes are interconnected if scientists know each other through working on joint projects or move from one organization to the other. As explained by the authors, technological overlap (in this case the number of technological classes in which two actors both hold patents) is meant as a necessary condition for cooperation as innovators share a minimum of common knowledge for understanding each other.

Given available data on metadata and technological complementarities, two perspectives of investigation could be pursued:

- a network of clean-tech start-ups and SMEs, where links between start-up$_i$ and start-up$_j$ result from tags used for both of them, which is based on a two-mode matrix X_c, where rows represent the companies and columns represent the tags;
- a network of clean-tech tags, where links between tag$_i$ and tag$_j$ result from the co-existence of tag$_i$ and tag$_j$ in the same start-up, which is based on a two-mode matrix X_t, where rows represent the tags and columns represent the companies.

In the former case, the square matrix indicating the number of links aij between start-up$_i$ and start-up$_j$ would be called the adjacency matrix Ac, which is computed as the product of X_c and its transposed (X'_c), while in the latter the square matrix indicating the number of links a_{ij} between tag$_i$ and tag$_j$ would be called the adjacency matrix A_t, which again is computed as the product of X_t and X'_t.

Cantner and Graf (2006) adopt the former perspective: nodes represent innovators and links between innovators are formed whenever they patent in the same technological class. For the purpose of identifying technological hotspots and complementarities the present work follows the latter perspective: nodes now represent products, markets and technologies (and not high-tech companies). Therefore, the proposed network analysis looks at links between tags (which are formed whenever these co-occur in the same company) and not between companies

(which instead would be formed whenever these are active in the same market, are specialized in the same R&D domain, offer the same product, make use of the same technology).

A network analysis on tags is the most suitable tool to look for complementarities between products, markets and technologies attracting most R&D efforts in an industry. To visualize data, the open-source network analysis software package Gephi is used (see Bastian et al. 2009).

3.2. The dataset

CrunchBase is the world's most comprehensive database on high-tech companies, accessible to everyone through an application-programming interface (API). Founded in 2007, CrunchBase began as a simple crowd sourced database to track high-tech start-ups covered on TechCrunch (one of the most highly regarded blogs concerning technological innovations on the Internet). Today CrunchBase collects information on more than 200,000 profiles of start-ups and other companies and is maintained by tens of thousands of contributors.

Companies listed in CrunchBase are active all over the world in several high-tech industries including analytics, automotive, bio-tech, clean-tech, consulting, e-commerce, education, enterprise, finance, games and video, hardware, health, hospitality, manufacturing, medical, mobile, nano-tech, network hosting, news, real estate, security, semi-conductor, software, transportation, travel and web. For most start-ups the database includes information such as the city of registration and operating offices, number of employees, category code, total money raised, number and timing of financing rounds and tags related to products, markets, technologies, and so on.

Data from CrunchBase is increasingly used in research (Block and Sandner, 2009, 2011; Xiang et al., 2012; Yuxian and Yuan, 2012; Adcock et al., 2013; Marra et al., 2014). Block and Sandner (2009) tested the quality of information in CrunchBase by comparing the funding data in the US tech industry with the industry statistics published by the National Venture Capital Association (NVCA, a trade association representing the US VC industry). They found that the number of deals in the CrunchBase figures amounts to about 97% of the internet-related deals as reported by NVCA and estimated a high and statistically significant Pearson correlation

coefficient between the time series of new deals related to the two sources (r=0.67; p<0.05).

Our dataset includes 74.310 recent start-ups, all companies founded between 2001 and 2013 that raised funding in dollars and for which available data concern company foundation, industry category code, metadata (tags on products, technologies and markets), number of employees, funding information and offices location. The adoption of such inclusion criteria is motivated, respectively, by our focus on recent and innovative firms, by the high level of incidence in CrunchBase of US tech start-ups compared to other countries and by the descriptive purpose of our study. Among the 74.310 companies, a subset of 30.824 start-ups was enucleated as representative for the SI environment within CrunchBase. Hereinafter, the original set of data is called "Crunch" and the subset is called "SI".

The identification of the SI subset was a long and not trivial exercise. First all metadata per start-up included in the Crunch dataset were collected and ordered, and duplicates were removed. Second, out of more than 140 thousands unique tags in the Crunch dataset, about 7 thousands tags were selected by hand and on a case by case approach in line with the schematization of social and innovative activities provided in the "Social Innovation Guide" elaborated by the EC (2013) and, more generally, the definition in Hubert (2010). Authors have performed this step independently. Afterwards, results were compared, checked and approved by two SI industry experts and, finally, put together. Within the resulting long list, for example, can be found tags such as social, education, community, health, food, crowd-sourcing, recruitment, learning, green, employment, kids, clean-tech, no profit, philanthropy, teaching, knowledge, eco-friendly, and so on.

Accordingly, topics close to our understanding of SI include demography (migration and ageing of the population), environmental trends (water, climate change, waste, sustainable supply chains and energy solutions in terms of renewable and smart innovation), new community trends (diversity and the new community providing it solutions, workplace innovation, incubation in the new concept of digital society), poverty-related trends (poverty, social exclusion and child poverty, cultural activities that reduce poverty and new way to diffuse and provide education), trends in health and well-being (health inequities, happiness and caring, welfare and new services), trends of ethical goods and services (fair trade, local production, organic

8

development, sustainable productions, social entrepreneurship, philanthropy, corporate social responsibility and social franchising).

## 4. An observational study of SI in the US high tech industries

Social and innovative activities in the US represent a fast-growing and fragmented phenomenon (Smith Milway 2014): according to Pacific Community Ventures (2013) only a fifth of social enterprises are larger than $2 millions in budget, just 8% employ more than a 100 people, and 60% were founded since 2006. For this reason, an illustration of our SI subset is insightful in several respects.

In the SI subset, together with well known best practices, less renowned social and innovative start-ups emerge, such as: a start-up promoting creativity and entrepreneurship to building livable and sustainable cities; an e-democracy platform that utilizes crowd-sourcing to collaborate on issues, determine the most effective solution and organize actions to see change through to implementation; a non-profit organization at the front of a new movement that believes children and families deserve the help of technology tools; a start-up that strives to be a well designed easy to digest blog in the world of M2M; a platform and active media for creating social innovation through crowd-funding, petitions and open democracy; a social innovation platform that enables organizations to solve business problems by crowd-sourcing from a global talent.

First of all, the trend in the number of start-ups by year of foundation is presented (Figure 1): SI start-ups rise from 588 units in 2001 to 2.571 units in 2013, exceeding the threshold of 4 thousands units in the period 2010-2012. The visible decline in 2013 is due to the delay with which the tech community updates CrunchBase: in all likelihood information about some couple of thousands of start-ups has been added after the date of our data extraction, 30th December 2013.

Looking at the trend of both SI and Crunch start-ups per year of foundation it can be noticed that their ratio between the two raises in fourteen years (2001-2013) from 29.5% to 44,2%, with its peak (47%) in 2012.

Figure 1: Number of start-ups in the Crunch and SI datasets per year of foundation
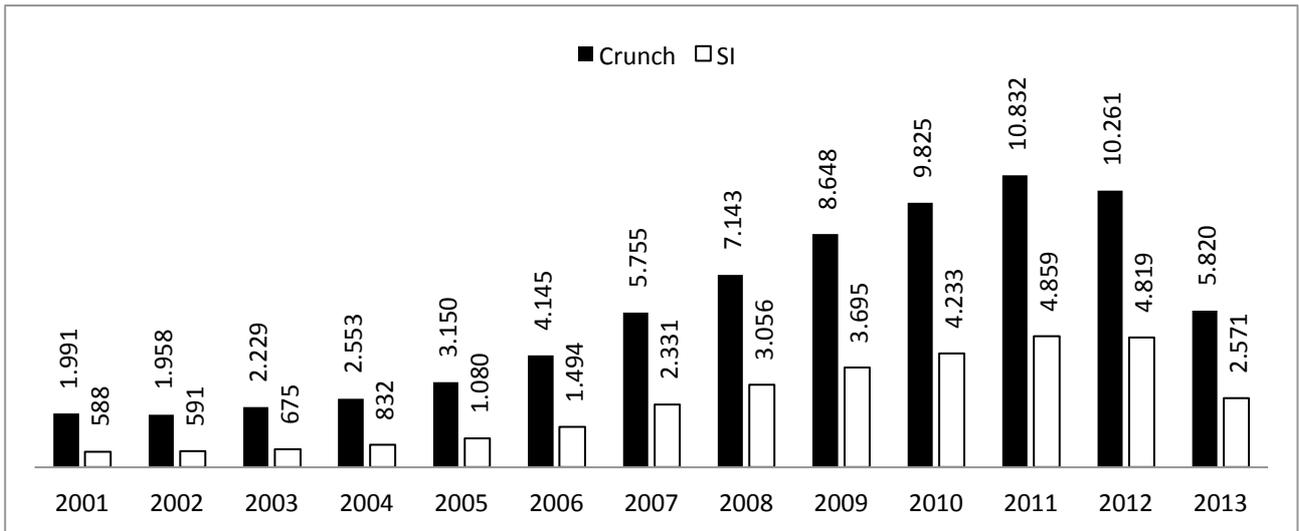
Source: CrunchBase

Second, the main traits of our dataset and subset are depicted by referring to the volume and annual trend in number of employees (Figure 2) and total money raised per year of foundation (Figure 3). The total number of employees in the SI subset is 621.440 vs. 1.357.727 in the Crunch dataset (45.7%), and money raised in the two datasets are respectively $ 82,6 billion vs. $ 241,8 billion (34.2% of the total amount referred to start-ups in our Crunch dataset).

Figure 2: Number of employees in the Crunch and SI datasets per year of foundation
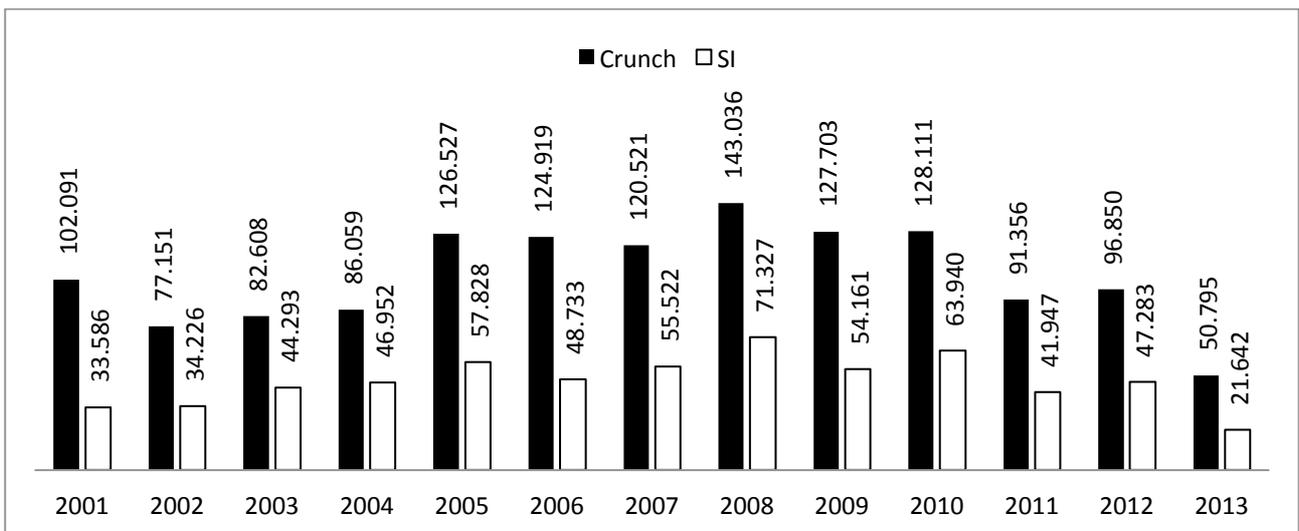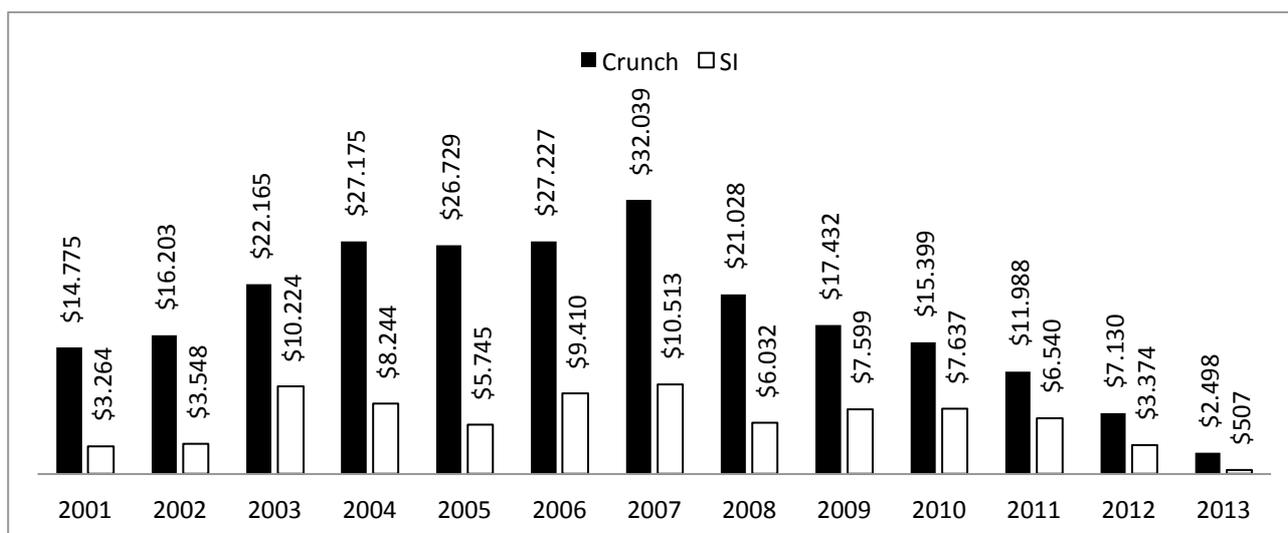

Source: CrunchBase

By comparing the number of employees per start-up in the two datasets emerges that over the observed time interval, on average, SI start-ups are always above Crunch ones, even if this gap has been reduced in the last 8 years. The decline in number of employees in the last years is explained by the fact that the count of employees is per year of foundation and not per year of hiring: the latter would have been more reliable information.

Looking at the money raised by start-ups per year of foundation it is interesting to notice the growing relevance of the SI subset within the Crunch dataset: the ratio between money raised by SI vs. Crunch start-ups increases between 2001 and 2012 from 22.1% to 47.3%, with its peak 54.6% in 2011. The decrease in the amount of money raised in the last few years is probably due to the shorter interval available for more recent start-ups to get funding compared to older ones.

Figure 3: Money raised (in million) by Crunch and SI start-ups per year of foundation

Source: CrunchBase

Data about the funding volume per year of investment is available (Figure 4) but lack of detailed information on date and amount for every single round does not allow to provide figures consistent with the above values: in particular, the total amount of money raised by start-ups per year of funding equals $ 60,8 billion vs. $ 82,6 billion (as provided above) for the SI subset (74%) and $ 179,8 billion vs. $ 241,8 billion (as provided above) for the Crunch dataset (74%). The above-mentioned data on funding is consistent with other estimates, at least since 2006, when social and innovative

activities gained real momentum and after CrunchBase was launched: for example, the Center for Venture Research (2011) at the University of New Hampshire reports that angel investments increased from $ 17 billion to $ 20 billion between 2009 and 2010. Taking for grant what Smith Milway (2014) sustains with regards to social enterprises, "*[In 2006] two global headlines raised the profile of social enterprise: Mohammed Yunus and the Grameen Bank won the Nobel Peace prize. And Bill Gates announced he was shifting his priorities from software development to social impact by moving full time to his foundation. In the broader U.S. population, two generations made clear that their interests, too, lay in helping society*" (Smith Milway 2014: 1).

Figure 4: Money raised (in million) by Crunch and SI start-ups per year of funding
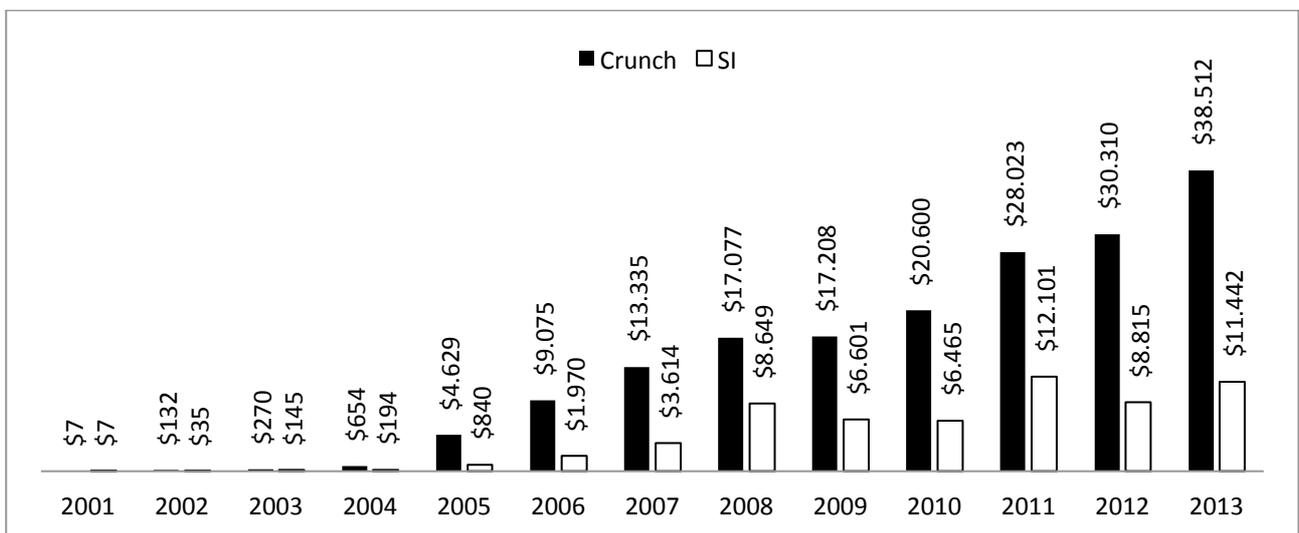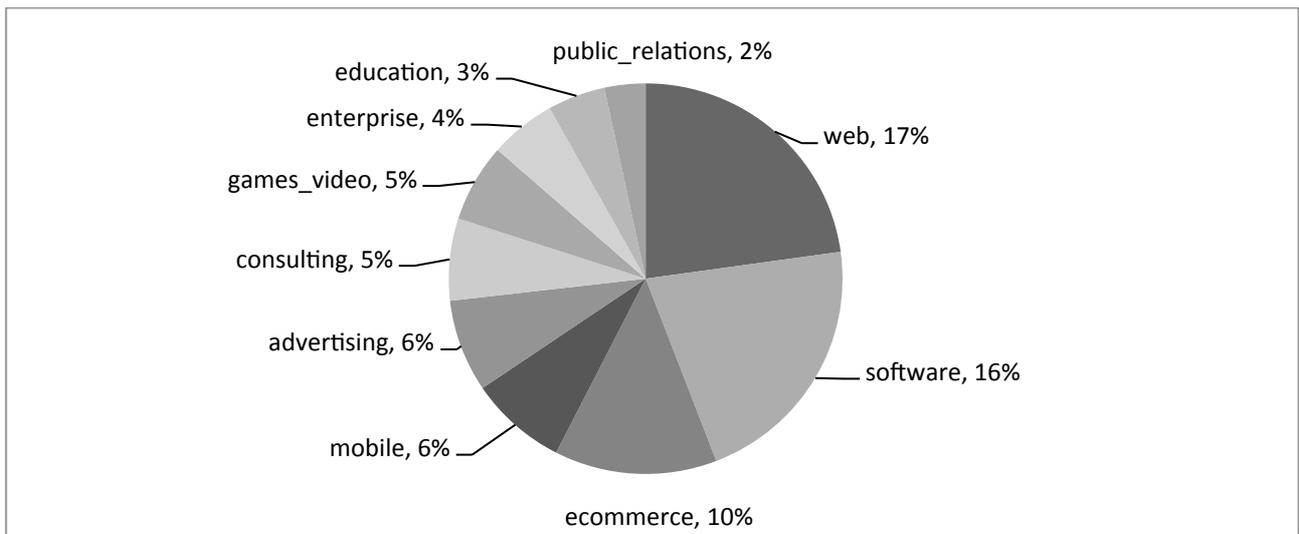

Source: CrunchBase

Again, looking at the trend of both SI and Crunch funding per year of investment since 2007 it can be noticed that their ratio is on average 36%, with peaks in 2008 (51%), 2009 (38%) and 2011 (43%).

Looking at the category code, there is no substantial variation in the industrial specialization in which SI and Crunch start-ups are active: this reinforces the fact that SI is a wide phenomenon that encompasses a large set of industries. The only difference emerges with the rise of the education sector among the first top ten industries: education is ranked 9th in the SI rank instead of biotech, which is 9th in the Crunch one (Figure 5).

Figure 5: First 15 industries ranked by number of start-ups in the SI dataset

Source: CrunchBase

This seems to confirm that SI is horizontal to markets, products and services, as implicitly recognized in literature (see, inter alia, Auerswald, 2009; Guida and Maiolini, 2013).

This circumstance induced to think about alternative ways to realize in which industry SI is more relevant. After a few attempts formulated on the basis of the BEPA schematization presented in Section 2, some text mining on tags resulted in a meaningful output. In CrunchBase a tag is a key word assigned to a single start-up, which takes in information on consumers and users' needs, products/services supplied, main functionalities, technological architecture and components, communication channels, programming languages, operating systems, supporting devices, and so on. This kind of metadata helps describe start-ups and allows finding them more easily. Tags, which are chosen informally and personally by every contributor to the database, represent a way to synthesize a lot of qualitative and descriptive information in something that can be collected, organized and graphically represented. By leveraging on information brought by such metadata, an in-depth investigation of the extent to which the US tech community is involved in social and innovative activities is provided.

First, sorting by relevance of tag in the SI subset, at the top of the ranking can be found terms such as mobile (2.261 times in the SI dataset), social-media (2.082), social-network (2.006), e-commerce (1.858), web-design (1.274), web-development

(1.027), big data (570), analytics (503), music (490), video (575), marketing (1.049), advertising (782), and so on. Again, this confirms that SI is widespread over different businesses. Nonetheless, in the top 100 tags there are a lot of industries, activities and concepts close to the BEPA understanding of SI: education (1.031 times in the SI dataset), community (719), health and healthcare (respectively, 474 and 352), food (349), news (339), entertainment (336), sharing (306), local (299), sports (269), training (255), crowd-funding (235), art (232), blog (226), college (222), recruitment (210), learning (200), etcetera. Aside from the mere count of times a tag recurs, the above metadata become more informative if put in relation to each other (see Section 5).

What about location of SI all over the world and, in particular, in US cities? Cities are often interpreted as the nurseries of innovative start-ups. Hoover and Vernon (1959) and Chinitz (1961) first put forth the "incubator hypothesis", by which small firms find it advantageous to locate initially at high density, central locations within the metropolis. This advantage is due to several factors, including access to specific knowledge, skilled workers, business opportunities, professional services. This hypothesis can be expanded in several ways and well fits the case of tech industries. Digital and innovative companies choose to locate differently from manufacturing ones: the former locate close to centers of research and science and to places where they have a good chance of rapid market penetration (Frenkel 2001; Marra et al., 2011).

Social and innovative start-ups have preference towards urban places where new ideas can spread and encounter raising users needs (Harris and Albury, 2009), information can be mediated and selected (Huston and Sakkab, 2006), markets can grow rapidly, entrepreneurial activities with no-profit origins can evolve and transform the underlying business model (Mulgan, 2006) and users can help in generating content (Cunningham et al., 2008).

Looking at the ratio between the number of SI and Crunch start-ups per city the following evidence raises (figure 6): Chicago is at the top of the list with 47.5% of SI start-ups in the Crunch dataset, followed by Berlin (45.3%), San Francisco (42.4%), New York (39.6%), London (39.3%), Austin (37%), Los Angeles (36.6%), Paris (36.2%), Palo Alto (35.8%), Mountain View (33.5%).

Figure 6: The ratio between the number of SI and Crunch start-ups per city
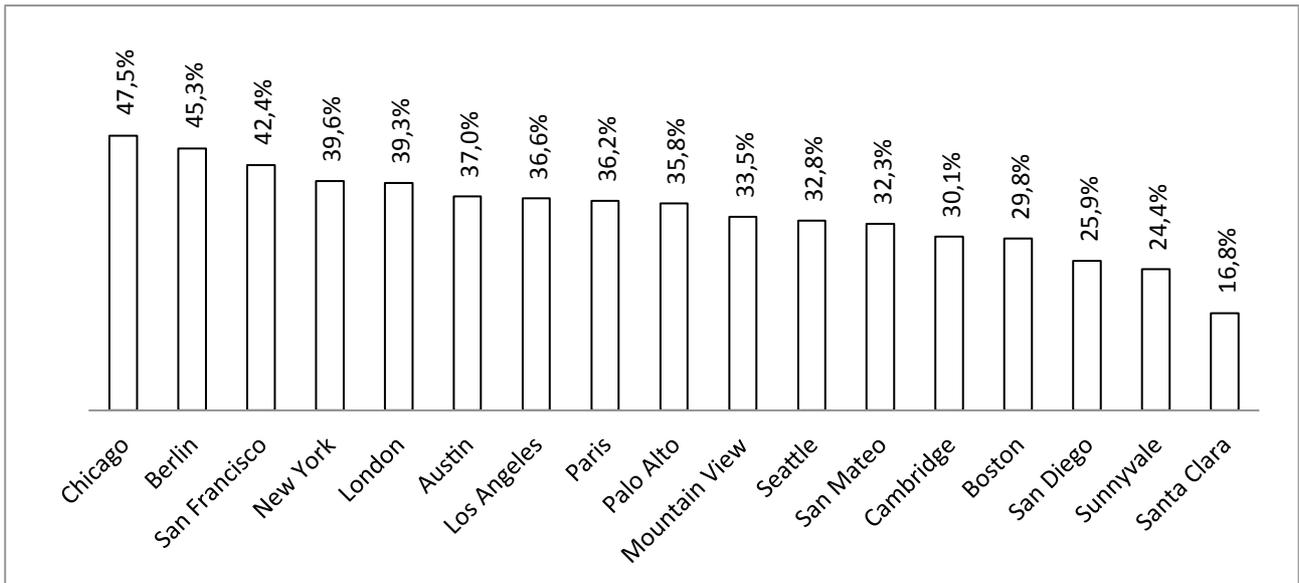
Source: CrunchBase

It is interesting to look at the SI industrial structure of cities to understand the underlying information flows from an economic point of view and to what extent industrial diversity facilitates the transmission of ideas (Jacobs 1969, 1984, Scherer 1982, Glaeser et al. 1992, Van Oort 2004).

All top ranked cities above present high percentage values in five industries: advertising, enterprise, mobile, software and web. Chicago tech industrial structure is rather diversified (7 industries presents percentage values above 5%) with high values in advertising (19%), analytics (5%), e-commerce (9%), enterprise (10%), mobile (5%), software (15%) and web (10%). In the last two sectors concentration is far below the levels in New York and San Francisco. New York presents a diversified economy, similar to Chicago with advertising (13%), e-commerce (6%), enterprise (5%), mobile (7%), software (16%) and web (17%) as relevant sectors. San Francisco, instead, presents a rather specialized urban economy: values above 5% are in social (6%), software (42%) and web (27%).

Understanding whether the urban environment fosters SI needs a targeted analysis of all possible determinants, including the existence and extent of investment funds and/or incubators specialized on SI. Moreover, metadata related to each start-up and tag connections between start-ups located in the same city give evidence about whether knowledge spillovers take place, in which sectors and to what extent and, finally, detect the underlying innovation trajectories.

15

5. A network analysis on metadata

As well known networks allow for a representation of a set of entities where each pair of entities is connected by links. Interconnected entities are represented by mathematical abstractions called nodes, and links connect each pair of nodes. Typically, a network is depicted in diagrammatic form as a set of dots or bubbles for nodes, joined by lines for edges.

In Figure 7 a preliminary graphical representation has been limited to the top 6 tags by count. The degree range has been set between 3.121 and 5.167, number of nodes and number of edges shown are, respectively, 6 and 15: weighting for the number of links between these tags (nodes), some robust relationships emerge between mobile and social, mobile and social-media, mobile and software, and social-media and social network.

Figure 7: A weighted network analysis between the top 6 SI tags
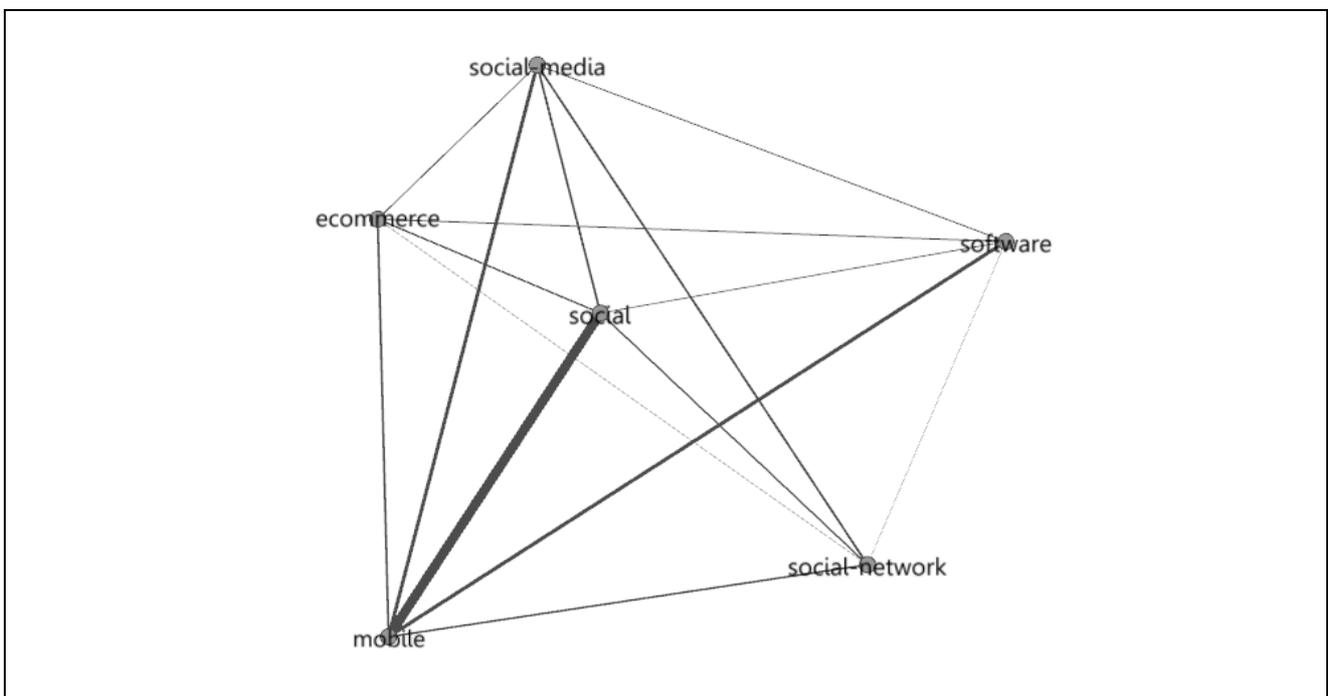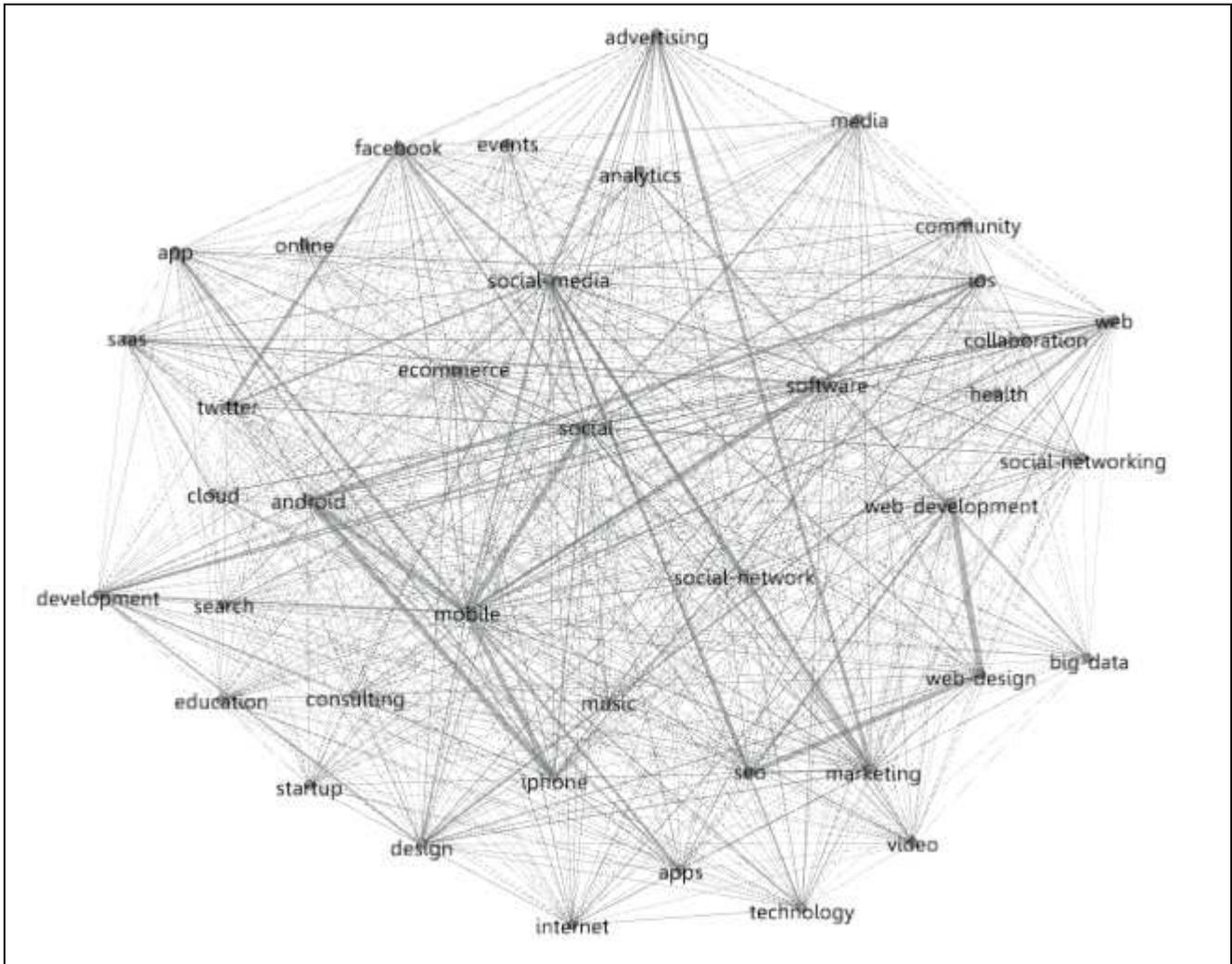

Source: CrunchBase

In Figure 8 the network between the above top tags and other tags in the SI dataset are shown. The graph between tags is strongly connected and presents the following

parameters: average degree equals 38.7, average weighted degree to 1433.1, graph density is 0.992, modularity is 0.195, network diameter equals 2 and average path length is set at 1.008.

Figure 8: A wider view on SI tags and connections

Source: CrunchBase

By using tools and plug-ins supplied by Gephi, it is useful to highlight non-trivial associations between tags and identify potential innovation trajectories of SI in US high-tech industries. First, social and innovative activities are associated to certain communication channels (mobile, web, internet, media), software development (web-design and development, apps, search engine optimization), operating systems (android, ios), specific devices (ipad, iphone), some major platforms (Facebook and

Twitter), technologies and services (cloud, saas, analytics) and industries (advertising, e-commerce, social-network, education).

These can be grouped in at least three sets of drivers of innovation: (a) mobile as device; (b) web and social as channels/platforms; (c) advertising, marketing, education and e-commerce as relevant industries. Evidence remarks that SI is "oblique", as defined by Pol and Ville (2009), and includes all types of new ideas and new kinds of social structures across markets, industrial sectors and technologies. Also the above interpretation seems to come with the vision of SI as a "vehicle" through which different waves of knowledge are merged to produce conditions of newness (see also Kanter, 1999). According to Cajaba-Santana (2013) the holistic interpretation of SI underlines the interaction between agents and social structures operating in different fields and using specific technologies as enablers of innovation (Eggers and McMillan, 2013). Also, emerge technological phases of the process that enable new forms of solution and open alternative ways to solve problems, according to the Kirzenarian model of innovation (Kirzner, 1999). Further vertical investigation could be carried out looking at every single node.

6. Conclusions and future research

SI is a fuzzy concept, difficult to define and circumscribe because of its complexity and novelty. In literature a consensus is emerging on the meaning of SI from a theoretical point of view, but a very few empirical studies exist, most of which are based on a "case by case" approach. Given this lack in the empirical literature the theoretical definitions presented in Section 1 served as a benchmark on which to build from a practical point of view our own understanding of SI, by investigating on a large-n base all social and innovative activities carried out by the US high-tech community.

Aside from figures on number of start-ups, employees, total funding raised, industries involved and most recurrent tags, a geo-referentiation of SI was provided by looking at some major cities. Urban areas are the arena where the collectivity identifies social challenges and try to devise new way to solve old problems or innovative and disruptive ones. The exchange of information between individuals with different experience and background is instrumental to encourage the emergence of new practices of SI. In this sense, cities preserve the potential for collective participation and facilitate the development of SI practices. Social and innovative companies have preference towards urban places where the innovation they bring can spread and

encounter raising users needs, information can be mediated and selected, product/service market can grow rapidly, given their very peculiar business model and users can help in generating content. Creation of content passes through the participation of individuals in a process of identification of issues and collective creation of solutions.

The definition of organic relationships between SI and suitable spaces can open up a new perspective for the SI road mapping, useful for planning and guiding the development of social and innovative activities (Phaal et al., 2004).

Observation of industry patterns suggested that SI is horizontal to markets and cannot be detected by filtering per industry. This fact induced to think about alternative ways to realize what SI means in practice. Some text mining on tags gave us the chance to derive insights on the cross-fertilization of SI: advertising, analytics, art, big data, blog, college, community, crowd-funding, e-commerce, education, entertainment, food, health and healthcare, learning, local, marketing, mobile, music, news, recruitment, sharing, social-media, social-network, sports, training, video, web-design and web-development.

Looking at connections between tags, some interesting insights were derived: social and innovative activities in US high-tech are associated to certain communication channels (mobile, web, media), software development (web-design and development, app), operating systems (android, ios), specific devices (ipad, iphone), some major platforms (Facebook and Twitter), several industries (advertising, gaming, e-commerce, social-network). In all likelihood this is motivated by the fact that one of the main driver of SI stays in the collective dimension. This view allowed to identify the underlying innovation trajectories of SI in US high-tech industries: mobile as device; web and social as channels/platforms; advertising, gaming and e-commerce as relevant industries.

The study aimed at bringing a contribution to the debate about SI and offering an original view of the phenomenon of SI to identify new trends and areas of interest. Next step in our research program will encompass investigation of tags and tags connections within the same industry and between start-ups located in the same city to understand which tags play the role of hot-spots for knowledge spillovers, where and to what extent and, finally, realize whether might exist different innovation trajectories at urban level.

7. References

- Adcock A.B., Lakkam M., and Meyer J. (2013). CS 224W Final Report Group 37, Stanford University.
- Auerswald, P. (2009). Creating social value. Stanford Social Innovation Review, 7(2), 51-55.
- Balconi, M., Breschi, S., Lissoni, F., 2004. Networks of inventors and the role of academia: an exploration of Italian patent data. Res. Policy 33, 127–145.
- Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.
- Bloch F. 2004. Group and network formation in industrial organization, In: Group Formation in Economics; Networks, Clubs and Coalitions. Demange G., Wooders M. (Eds.). Cambridge, UK: Cambridge Univ. Press.
- Block, J.H. and Sandner, P.G. (2009), What is the Effect of the Financial Crisis on Venture Capital Financing? Empirical Evidence from US Internet Start-Ups. Venture Capital - An International Journal of Entrepreneurial Finance, 11, 4, 295-309.
- Block, J.H. and Sandner, P.G. (2011), Venture capital funding in the middle of the year 2011: are we back to pre-crisis boom levels?. Strategic Change, 20: 161–169.
- Breschi, S., Lissoni, F., 2003. Mobility and Social Networks: Localised Knowledge Spillovers. CESPRI Working Paper No. 142.
- Bulut, C., Eren, H., & Halac, D. S. (2013a). Social innovation and psychometric analysis. Procedia-Social and Behavioral Sciences, 82, 122-130.
- Bulut, C., Eren, H., & Halac, D. S. (2013b). Which One Triggers the Other? Technological or Social Innovation. Creativity Research Journal, 25(4).
- Cajaiba-Santana, G. (2014). Social innovation: Moving the field forward. A conceptual framework. Technological Forecasting and Social Change, 82, 42-51.
- Cantner, U., Graf, H., 2006. The network of innovators in Jena: an application of social network analysis. Res. Policy 35, 463–480.
- Center for Venture Research (2011). The angel investor market in 2010: a market on the rebound (available at wsbe.unh.edu).
- Chinitz, B.J. (1961). Contrasts in Agglomeration: New York and Pittsburgh, American Economic Review, 51, 279-89.
- Eggers, W. D., & Macmillan, P. (2013). The Solution Revolution: How Business, Government, and Social Enterprises are Teaming Up to Solve Society's Toughest Problems. Harvard Business Review Press.
- Frenkel, A. (2001). Why High-technology Firms Choose to Locate in or near Metropolitan Areas. Urban Studies (Routledge), 38(7), 1083-1102.
- Glaeser E.L., Kallal H., Scheinkman J. and Shleifer A. (1992). Growth in cities, Journal of Political Economy, 100, 1126-1152.
- Guida M.F. and Maiolini R. (2013), Il Fattore C per l'Innovazione Sociale, Rubbettino, Soveria Mannelli (CZ).
- Hamalainen, T. J. and Heiskala, R. (2007), Social Innovations, Institutional Change and Economic Performance. Edward Elgar, Cheltenham, pp. 52-79.

- Harris, M., & Albury, D. (2009). The innovation imperative. Why Radical Innovation is Needed to Reinvent Public Services for the Recession and Beyond, Londres, National Endowment for Science, Technology and Arts (NESTA).
- Hoover E.M. and Vernon R. (1959) Anatomy of a Metropolis. Cambridge, MA: Harvard University Press.
- Howaldt J. and Schwarz, M. (2010). Social Innovation: Concepts, research fields and international trends. K. Henning, & F. Hees (Eds.). IMA/ZLW.
- Hubert, A. (2010), Empowering people, driving change: Social innovation in the European Union, Bureau of European Policy Advisors (BEPA). Available.
- Huston, L., Sakkab, N. (2006). Connect and develop. Harvard business review, 84(3), 58-66.
- Jackson, M., 2006. The Economics of Social Networks, In: The Econometric Society, Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society. Blundell, R., Newey, W., Persson, T. (Eds.). Cambridge, UK: Cambridge Univ. Press.
- Jackson, M., 2011. An Overview of Social Networks and Economic Applications, In: Social Economics. Benhabib, J., Bisin, A., Jackson, M.O. (Eds.). Amsterdam: Elsevier.
- Jacobs J. (1969). The Economy of Cities. Random House, New York.
- Jacobs J. (1984). Cities and the Wealth of Nations: Principles of Economic Life. Random House, New York.
- Kanter, R. M. (1999). From spare change to real change: The social sector as beta site for business innovation. Harvard business review, 77, 122-133.
- Kirwan, J., Ilbery, B., Maye, D., & Carey, J. (2013). Grassroots social innovations and food localisation: an investigation of the Local Food programme in England. Global Environmental Change, 23(5), 830-837.
- Kirzner, I. M. (1999). Creativity and/or alertness: A reconsideration of the Schumpeterian entrepreneur. The Review of Austrian Economics, 11(1), 5-17.
- Marcy, R. T., Mumford, M. D. (2007). Social innovation: Enhancing creative performance through causal analysis. Creativity Research Journal, 19(2-3), 123-140.
- Marra A., Carlei V., Crociata A. (2011) Economic diversity and employment levels in Italian provinces over the period 1951-2001: Does a recurrent industrial pattern exist? Modern Economy, 2, 3, 438- 445.
- Marra, A., Luciani, M., Antonelli, P., 2014. Data analytics nel comparto high-tech statunitense: impatti, evoluzione e tutela della privacy. L'industria 35, 319-338.
- Maruyama, Y., Nishikido, M., & Iida, T. (2007). The rise of community wind power in Japan: Enhanced acceptance through social innovation. Energy Policy, 35(5), 2761-2769.
- Mulgan G., Ali R., Halkett, R., Sanders, B., (2007): In and out of sync. The challenge of growing social innovations. Research report. London.
- Mulgan, G. (2006). The Process of Social Innovation. Innovations: Technology, Governance, Globalization, 1(2), 145-162.
- Mumford, M. D. (2002). Social innovation: Ten cases from Benjamin Franklin. Creativity Research Journal, 14, 253–266.

- Mumford, M. D., & Moertl, P. (2003). Cases of social innovation: Lessons from two innovations in the 20th century. Creativity Research Journal, 15(2-3), 261-266.
- Murray R., Mulgan G., Caulier-Grice J., (2008): How to Innovate: The tools for social innovation.
- Owen-Smith, J., Riccaboni, M., Pammolli, F., Powell, W.W., 2002. A comparison of U.S. and European university-industry relations in the life sciences. Management Science 48, 24–43.
- Phaal, R., Farrukh, C. J., & Probert, D. R. (2004). Technology roadmapping—a planning framework for evolution and revolution. Technological forecasting and social change, 71(1), 5-26.
- Phillips, F. (2011). The state of technological and social change: Impressions. Technological Forecasting and Social Change, 78(6), 1072-1078.
- Phills, J. A., Deiglmeier, K., & Miller, D. T. (2008). Rediscovering social innovation. Stanford Social Innovation Review, 6(4), 34-43.
- Pol, E., Ville, S. (2009). Social innovation: Buzz word or enduring term?. The Journal of Socio-Economics, 38(6), 878-885.
- Rodima-Taylor, D. (2012). Social innovation and climate adaptation: Local collective action in diversifying Tanzania. Applied Geography, 33, 128-134.
- Scherer F.M. (1982). Inter-industry technology flows in the United States, Research Policy, I, 227-45.
- Van OOrt F.G. (2004). Urban Growth and Innovation. Spatially Bounded Externalities in the Netherlands. Ashgate, Aldershot.
- Xiang G., Zheng Z., Wen M., Hong J., Rose C., and Liu C. (2012). A Supervised Approach to Predict Company Acquisition with Factual and Topic Features Using Profiles and News Articles on TechCrunch. ICWSM 2012.
- Yuxian E.L., Yuan, S.T.D. (2012). Where's the Money? The Social Behavior of Investors in Facebook's Small World." 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.