

# **LISTENING TO THE BUZZ: EXPLORING THE LINK BETWEEN FIRM CREATION AND REGIONAL RELATIONAL STRUCTURES AS REFLECTED BY SOCIAL MEDIA**

CARLO CORRADINI

University of Birmingham, Edgbaston, Birmingham, UK.

EMMA FOLMER

University of Groningen, Netherlands.

ANNA REBMANN

King's College London, UK.

## **INTRODUCTION**

In this paper, we explore how to expand our knowledge of informal relational structures and how they influence entrepreneurship by analysing big data to listen to the ‘buzz’ – the vibrancy and knowledge exchange propensity of localised informal communication flows. In the empirical literature, buzz has been treated as a synonym for density of interaction and it remains subsumed within the concept of agglomeration economies and the density of economic activity (Qian et al. 2013). However, this ignores an important dimension of the concept – the atmosphere and content of the buzz, reflecting “nondeliberate knowledge-exchange and information-exchange propensities” (Ansheim et al. 2007: 659). Such propensity has usually been assumed to be a function of the knowledge and human capital available in a region. Yet, in line with research on the importance of the place-specific informal institutional and social context in shaping relational structures (Gertler, 2003; Saxenian, 1996; Storper, 1997), we argue that for the same levels of knowledge and interaction, some places may be defined by a more innovative atmosphere supporting experimentation with new ideas and innovative practices than others. Looking at the content of buzz would allow us to better understand the nature of informal knowledge exchanged within localised communities and test the seminal observation by Marshall that knowledge spillovers are “as it were in the air” (Marshall, 1890, p.198). Until now, it has not been possible to capture the intangible nature of the discourse in the buzz and test whether regions with buzz that represents a more innovative atmosphere experience higher rates of entrepreneurship. We propose a novel approach to measure the innovative atmosphere which is an intangible element of learning regions, reminiscent of the ‘industrial atmosphere’ discussed by Marshall, within the framework of the knowledge spillover theory of entrepreneurship. Using all geo-tagged Tweets across 168 NUTS3 regions in the UK for the year 2014, our results offer evidence, robust to different model specifications, that regions characterised by a relatively higher intensity of discussion and vibrancy around topics related to innovation may provide a more effective set of informal relational resources for sharing and recombination of ideas, defining regional capabilities to support and facilitate entrepreneurial processes.

## **FRAMEWORK AND HYPOTHESES**

### **Spatial proximity, knowledge flows and entrepreneurship**

Economic geography has long studied contextual factors of economic activity and offers useful insights for the analysis of entrepreneurship. Building on the seminal insights on territorial agglomerations by Marshall (1890), the literature has described regions as loci for interactive learning underpinning innovation and new knowledge creation (Bathelt and Glückler 2003; Camagni 1991; Storper 1997). This perspective has also served as a framework of reference for the analysis of entrepreneurial processes at the spatial level. In particular, the knowledge spillover theory of entrepreneurship (Acs et al., 2009; Audretsch, 1995) posits that knowledge created but left unexploited by incumbent actors may be identified, valued and recombined with other ideas by other spatially embedded agents. Thus, “contexts rich in knowledge will generate more entrepreneurial opportunities” (Audretsch & Keilbach 2008), with human capital defining regional absorptive capacity to recognise and exploit such opportunities leading to the creation of new companies (Audretsch & Lehmann, 2005; Qian & Acs, 2013; Qian et al. 2013).

These dynamics present a strong spatial dimension as the exchange of tacit and uncodified knowledge (Polanyi 1967) underpinning entrepreneurial opportunities occurs more effectively through spatially embedded interactions reflecting the localised nature of knowledge spillovers (Jaffe et al., 1993; Sonn and Storper, 2008). This is notably reflected in the concept of “buzz”, defining the complex web of information and knowledge exchange ecology created through co-location (Bathelt et al., 2004; Storper and Venables, 2004).

### **Listening to the buzz**

Of places with similar levels of knowledge creation and interaction, some may be defined by a stronger drive to (re)discuss ideas and search for new directions. Bathelt et al. (2004: 38) have defined ‘buzz’ as “the information and communication ecology created by face-to-face contacts, co-presence and co-location of people and firms within the same industry and place or region”. Scholars thus have traditionally underlined spatially bounded interaction as being critical for buzz (Bathelt et al., 2004; Gertler, 2003; Storper and Venables, 2004). Yet, buzz is not simply co-location for face to face interaction, but it also entails socialisation effects and psychological motivation (Storper and Venables, 2004). Asheim et al. (2007) offer a definition following the idea of ‘noise’ or ‘local broadcasting’ as put forward by Grabher (2002), and state that “buzz refers to *nondeliberate* knowledge-exchange and information-exchange propensities” (Asheim et al. 2007: 658) which occurs mainly outside of formal collaboration.

In this sense, buzz can be conceptualised broader than tacit knowledge transfer through personal interaction. Buzz also has a community dimension, as it is also shaped by “shared cultural traditions and habits [, with] actors continuously contributing and benefiting from the diffusion of information, gossip and news by just ‘being there’” Bathelt et al. (2004: 38). In this sense, it reflects the localised set of norms and practices akin to the seminal concept of industrial atmosphere discussed by Marshall (1890), pointing to the importance of the circulation of knowledge – such as trade secrets and know-how - to firms in a particular region or place as if it were in the air. Liking buzz to industrial atmosphere therefore implies that buzz does not simply refer to high interaction, but it also includes a localised set of norms and practices which may differ from place to place (Bathelt et al. 2004).

Thus, having established that buzz is a crucial aspect of co-location of economic activity, the actual content of buzz may become an important aspect of knowledge creation and the diffusion of entrepreneurial opportunities. Knowledge-intensive, volatile and unstandardized economic activities especially benefit from a production climate that is characterised by a

common conceptual framework that endorses experimentation and diversity of ideas (Scott, 1997). Entrepreneurial activity and innovation thrive in an environment with a certain attitude of economic dynamism encouraging the development of new products and services. Such economic dynamism has traditionally been captured by analysing local levels of human capital and occupational structure. While these factors are undoubtedly important for such economic dynamism, previous research has underlined the relevance of other aspects of localised communities. Building on the arguments by Florida, several authors show that higher proportions of people in creative industries and ‘bohemians’ are linked to regional and city-level firm entry (Audretsch and Belitski 2013; Lee et al. 2004;). This view has been criticised along two main elements. First, some authors argue the creative capital of regions is simply a proxy for human capital (Glaeser, 2004). Others have pointed out that measuring creativity using creative occupations may lead to excluding many other jobs that are not classified as creative but indeed present significant aspects of innovative content (Boschma and Fritsch, 2009). Levels of creativity and innovation in a region are thus important but may not be accurately captured by measures of human capital and occupational structure.

In a similar vein, some authors (Storper and Allen 2009; Storper et al. 2015) have criticised existing approaches for being centred on the mere presence of people with creative and innovative capabilities. Instead, this stream of research argues it is place-specific norms and values in local communities that define the ‘industrial atmosphere’ of places. Consequently, the propensity to exchange ideas goes beyond the level of human capital available and it is not simply defined by creative people moving in. Conversely, places may offer a nurturing innovative environment through their informal norms. This is another aspect of buzz, which also captures peer pressure and may reveal the presence of an innovative atmosphere due to the propensity to constantly exchange ideas.

### **Buzz and entrepreneurship**

In line with these arguments, we put forward that analysing local discursive practices is a novel way of understanding entrepreneurial climates. A local buzz with an innovative atmosphere exposes potential entrepreneurs to valuable knowledge and a continuous re-discussion of information about opportunities and encourages them to take entrepreneurial action. Greater intensity of buzz around innovation in a region, for similar levels of density and knowledge creation, mean that more people are continuously contributing to the diffusion of information and news. Thus, we propose:

*Hypothesis 1: more regional buzz on innovation will be positively related to regional rates of entrepreneurship.*

## **DATA AND METHODS**

### **Measuring buzz**

We offer a novel approach to allow for the quantification of buzz by proposing that data from Twitter can be used as a proxy for buzz within a region. We use all geo-tagged UK tweets from 2014, providing the exact location in terms of longitude and latitude of the person tweeting at the time of the tweet. Twitter only allows for original content to be geo-tagged; re-tweets are

not geo-tagged and are not used in this analysis (Sloan and Morgan, 2015). This means we can identify the regions in which people are discussing entrepreneurship. In particular, we used aggregated information across NUTS3 regions based on over one million tweets per day across the UK in 2014. This is then merged with socio-economic data from the UK Office for National Statistics (ONS) as discussed below.

An issue inherent in using big data such as twitter is its representativeness. Research has shown that UK twitter users do differ from the general population, as Twitter users are younger and there is a slight under-representation of users in routine occupations, whilst gender distribution is the same as the UK census (Sloan 2017). Similarly, considering differences between those who use twitter and the sub-section of tweeters who use geo-location when tweeting, Sloan and Morgan (2015) point to differences around 1% for gender, age and occupation category. Crucially, whilst Twitter data are not representative of the general UK population and geo-tagged tweets are not fully representative of Twitter users, there is no indication that these differences in characteristics of who tweets or geo-tags tweets varies over geographic space. There are slight differences in tweeting patterns in rural areas and cities (Gross et al. 2017), mostly in terms of frequency of tweets, but not in terms of content or the population tweeting. Thus, this subset of people selecting into Twitter and geo-tagging can be assumed to be consistent across regions.

With respect to the connection between the content of Tweets and their location, a stream of recent research of online social networks has shown that these are connected to place and are important as local forms of communication (Gross et al. 2017; Stephens & Poorthuis, 2015). Grabher (2002) argues that online interaction does not replace offline interaction but can reflect it or even ‘multiply’ it. Similarly, Tranos and Nijkamp (2013) explore the relationship between physical distance and virtual geography, highlighting the spatiality of the Internet. Accordingly, Twitter data have been used already to provide robust evidence and accurate measures on the spatial variation of a variety of topics, ranging from rates of heart-disease (Eichstaedt et al., 2015), unemployment (Llorente et al., 2015) to dominant demographic dimensions (Bokányi et al., 2016). This supports the use Twitter data to develop a proxy measure of buzz in a region.

## **Variables and estimation approach**

An exploratory econometric model is proposed where the regional rate of firm entry is regressed on our measure of buzz, controlling for other regional determinants of entrepreneurship to identify whether the vibrancy around trending twitter keywords associated with innovation and entrepreneurship presents a positive and significant correlation with entrepreneurship, *ceteris paribus*, using data from 168 UK regions.

In line with previous research (Corradini and De Propris 2015; Lee et al. 2004), we measure entrepreneurship looking at the rate of new firms over regional population. To listen to buzz, we propose that the content that Twitter users exchange can be used as proxy of spatially-bounded informal information flows. We analyse the content of tweets in each region, looking for the relative frequency of the word “innovation” (number of times “innovation” present in a tweet over the total number of tweets in thousands for each NUTS3 region) creating the variable INNOBUZZ.

To control for potential confounding factors, we also include in the model a set of variables that reflect important determinants of firm birth at the regional level, as identified by previous literature (Armington and Acs 2002; Fotopoulos 2013). First, we include GDP per

capita in purchasing power standards and GDP growth reflecting the level of economic development and increases in local demand. To capture knowledge creation and absorptive capacity, we include human capital defined as the share of the working population (people aged 25-64) in the region with tertiary education, which is equivalent to at least degree-level qualifications, R&D investment to reflect input into the knowledge production function for the region, and finally we add a creative index defined by the share of people in creative industries (defined by the ONS) over total employment (Lee et al., 2004; Sleuwaegen and Boiardi, 2014). To proxy intensity of interaction, we use the established approach in the literature based on population density, measured as the number of people per square kilometre at the NUTS3 level. We also test an alternative measure reflecting density of interaction in pooled labour markets, based on the number of firms over total population (Armington and Acs, 2002). Finally, we include a proxy for sectoral diversification calculated as the inverse of the Herfindahl index at the 4-digit level.

In the analysis, we first estimate our model using OLS regression analysis with cluster robust standard errors to control for heteroscedasticity. For robustness, we also try to apply two-stage instrumental variable (IV) regression to control for potential reverse causality in our model and, more importantly, potential omitted variable bias. This allows us to explore another potential advantage of using data from social media, as we identify two potential instruments in the relative number of tweets containing the words “investor” and “conference”. Both turn out to present the ideal characteristics for this analysis as they are not significant predictors of firm entry yet are strongly correlated to INNOBUZZ once the other independent variables are controlled for. However, we note the limited sample size may not allow for robust IV estimates and a longitudinal dataset may ultimately be necessary to shed a more conclusive light on this.

## RESULTS

Exploring our key variable INNOBUZZ, we find a strongly uneven distribution across the various NUTS3 regions in the UK. This variable is higher in areas defined by high population density and high GDP per capita, both elements characterising larger cities where buzz is expected to be more prominent. When looking at absolute values for INNOBUZZ, they are particularly high for the largest cities, starting from central London, but also Manchester, Birmingham, Nottingham as well as Belfast. The regions of Oxfordshire and Cambridgeshire are similarly included in the top decile. Thus, we find that the distribution of INNOBUZZ across UK regions is not random, reflecting the agglomerated nature of economic activity. INNOBUZZ also reveals a strong correlation with measures of GDP and density, in line with the expected connection of an innovative atmosphere with more developed areas and places with strong interaction.

The results from the regression analysis can disentangle these different effects. Across different model specifications starting from a baseline form including only GDP and education as controls, up to the full model with all control variables included, we find a positive and statistically significant effect for INNOBUZZ. This provides evidence for our hypothesis of an important effect of an innovative atmosphere on entrepreneurial activity, even after controlling for socio-economic characteristics of the area reflecting established determinants of firm entry such as GDP, knowledge and human capital and density of interaction. Considering IV regression, F-statistics of instrumental variables in the first stage are consistently larger than 10,

suggesting instruments are not weak. We also run a test of overidentifying restrictions, pointing to the validity of the instruments. The main findings are robust to the IV regression specification.

## CONCLUSIONS

In this paper, we have used Twitter data to develop a quantitative measure of buzz. This constitutes a significant methodological contribution as whilst there have been qualitative studies of buzz, there is (to our knowledge) currently no quantitative measure for buzz. Using Twitter data to capture buzz also represents an innovation in how to use social media in economic geography as well as entrepreneurship research, where almost all papers have been focusing on how entrepreneurs are using twitter rather than what macro-level trends we can discern from the data.

Having developed this quantitative measurement of buzz allows us to contribute further to the literature by enabling us to test if and how localised informal knowledge exchange benefits entrepreneurship and innovation. We find that regions which are characterised by higher levels of “buzz” representing an innovative atmosphere – more discussion and vibrancy about topics related to innovation – do have higher levels of entrepreneurship. If we take the stance that controls on R&D expenditure, education levels and relative weights of creative occupations are able to capture the level of human capital and knowledge creation in a region, then our results provide strong evidence in favour of a critical role of informal institutions, and in particular the presence of an innovative atmosphere consisting of the exchange of ideas and knowledge conducive to the diffusion and exploitation of entrepreneurial opportunities. This research thus contributes to our understanding of regional entrepreneurial dynamics and the contextual factors that foster entrepreneurial regions. Our study provides evidence that in conjunction with interaction and the presence of human capital, it is thus important to look at the *content* of information exchange as source of entrepreneurial opportunities. If we take the position that some aspects of knowledge and human capital are not properly captured by these variables, then our results indicate data from social media could be used to capture intangible aspects related to skills which are currently omitted from previous research. This provides quantitative evidence to underpin the theories of buzz as an informal relational resource which acts a further channel for knowledge spillovers that allow entrepreneurs to discover and exploit opportunities.

The paper has followed an exploratory approach to provide empirical evidence of the content of buzz. As such, the results from our analysis must be interpreted considering some possible limitations. In particular, the limitations on the longitudinal dimension of our dataset do not allow us to control for unobserved heterogeneity across regions to offer further robustness to our results. Further research is currently undergoing to generate a panel dataset. Also, our approach for measuring an innovative atmosphere presents advantages but also limitations that clearly define opportunities for further research. On one side, the simplicity of our approach allows for a straightforward application for innovation policy. At the same time, more complex algorithms and machine learning could be implemented to identify a wider set of keywords for capturing innovative atmosphere and develop a more nuanced indicator.

## REFERENCES AVAILABLE FROM THE AUTHOR