How do cognitive proximity and knowledge networks affect firms' innovation? Evidence from micro and small firms in Johannesburg, South Africa

Jefferson Galetti and Fiona Tregenna

SARChI Industrial Development Working Paper Series

WP 2022-04

March 2022

How do cognitive proximity and knowledge networks affect firms' innovation? Evidence from micro and small firms in Johannesburg, South Africa

DSI/NRF SOUTH AFRICAN RESEARCH CHAIR IN INDUSTRIAL DEVELOPMENT

Jefferson Galetti and Fiona Tregenna

SARChI Industrial Development Working Paper Series

WP 2022-04

ISBN 978-1-77630-397-7

March 2022

About the South African Research Chair in Industrial Development (SARChI-ID)

The DSI/NRF South African Research Chair in Industrial Development conducts research, builds capacity and undertakes public and policy engagement in the field of industrial development. Activities focus on research projects; training and supervision of graduate students; hosting postdoctoral fellows and research visitors; and various projects, often in conjunction with partners, such as conferences, workshops, seminars, training courses, and public and policy engagements. SARChI Industrial Development is hosted at the University of Johannesburg, where it operates as a centre located in the College of Business and Economics.

Funding acknowledgement

The South African Research Chairs Initiative (SARChI) was established in 2006 by the then Department of Science and Technology (DST), now known as the Department of Science and Innovation (DSI), and the National Research Foundation (NRF). The Chairs are designed to attract and retain excellence in research and innovation at South African public universities. The funding support of the DSI and the NRF through Grant Number 98627 and Grant Number 110691 for the South African Research Chair in Industrial Development has made this working paper series possible.

Recommended citation

Galetti, J., and Tregenna, F. (2022). How do cognitive proximity and knowledge networks affect firms' innovation? Evidence from micro and small firms in Johannesburg, South Africa. SARChI Industrial Development Working Paper Series WP 2022-04. SARChI Industrial Development, University of Johannesburg.

Disclaimer

Working Papers can be downloaded from https://www.uj.ac.za/faculties/college-of-business-and-economics/schools/school-of-management/south-african-research-chair-in-industrial-development/working-paper-series/ in PDF (Adobe Acrobat) format.

Abstract

Cognitive proximity between firms, and their embeddedness in a knowledge network, can affect firms' innovation performance. While some studies have begun analysing this in developing countries, there is a dearth of evidence in Africa, and for small and especially micro enterprises and informal enterprises more generally. To fill this gap and to assess whether cognitive proximity and knowledge network centrality matter for firms' innovation performance, we utilise rich new survey data covering 711 micro and small manufacturing firms in Johannesburg, South Africa. We develop and apply a new multidimensional measure of cognitive proximity, using the information on skills most needed in firms' activities, their technological relatedness, and the types of external co-operation in which they engage. We find that the relationship between cognitive proximity and innovation follows an inverse U-shaped curve associated with the 'proximity paradox', in which proximity enhances the probability of innovating, but too much proximity may lead to a cognitive lock-in. Firms' network position, measured using both degree centrality and betweenness centrality, also affects innovation outcomes. In an extension, we find interesting differences in how both cognitive proximity and firms' network positions affect the degree of novelty in innovation outcomes.

Keywords: innovation, micro and small enterprises (MSEs), cognitive proximity, knowledge networks, South Africa

JEL codes: O33, R11, D85, L26

About the Authors

Jefferson Galetti, Whitaker Institute for Innovation and Societal Change, National University of Ireland Galway and DSI/NRF South African Research Chair in Industrial Development, University of Johannesburg. Email: jefferson.galetti@nuigalway.ie.

Fiona Tregenna, DSI/NRF South African Research Chair in Industrial Development and Professor of Economics, University of Johannesburg. Email: ftregenna@uj.ac.za.

Acknowledgement

Support for this research was received under the project 'Community of Practice in Industrialisation and Innovation' (grant number 110691), hosted by the DSI/NRF South African Research Chair in Industrial Development (grant number 98627), University of Johannesburg.

Table of Contents

List of Tables

List of Figures

1. Introduction

Knowledge flows across firms can be important for learning and innovation, and consequently for firms' performance. The concept of distance is usually associated with interactive learning, as knowledge spillovers are assumed to occur mainly between firms in the same location (McCann, 2007). While geographical proximity has long been shown to be effective in transmitting knowledge across firms (Bell & Zaheer, 2007), there is growing recognition that knowledge externalities are not 'in the air' and are not automatically available for all firms located in the same place (Boschma & Ter Wal, 2007; Fitjar & Rodríguez-Pose, 2017). Alternative dimensions of relational proximity based on some logic of similarity between agents (Boschma, 2005; Torre & Rallet, 2005) also play a role in reducing uncertainty, solving coordination problems, facilitating interaction, and enhancing learning and innovation (Capone & Lazzeretti, 2018; Fitjar et al., 2016). Cognitive proximity refers to similarities in levels of knowledge and expertise among firms, which potentially allow them to exchange effective information and learn from each other (Boschma, 2005; Nooteboom, 2000).

Recent empirical studies have identified cognitive proximity as a driving force in establishing informal networks in which knowledge circulates (Balland et al., 2016; Capone & Lazzeretti, 2018). These networks are social constructs, where cognitive proximity on the one hand brings similar firms together in the densest areas, thus facilitating knowledge exchange, while on the other hand, it pushes less cognitively close firms to peripheral regions with low levels of knowledge flows (Gertler, 2003; Hidalgo et al., 2007). Thus, knowledge is not freely available to all firms in the same location; rather, access to knowledge can be conditional on the degree of cognitive proximity between firms and their position in the network. In this context, the connections and embeddedness of more central firms in the network tend to increase their likelihood of accessing relevant knowledge flows, learning more from the local 'buzz' and enhancing innovation performance (Boschma & Ter Wal, 2007; Gilsing et al., 2008; Stuck et al., 2016).

Despite the growing interest of scholars in cognitive proximity and innovation, there are some research gaps. One area in which a gap exists is the regional aspect. Relationships within a network tend to be context-specific and to vary depending on local capabilities and firms' characteristics (Gertler, 2003). Regional disparities might influence the context in which firms are embedded and how they exchange knowledge and innovate. In particular, we can identify three key aspects of this location factor that require further research for an improved understanding of how localised networks relate to innovation.

First, firms and workers in a particular geographical area tend to share common characteristics such as language, norms and legal frameworks, which generate and reinforce local trust (Bell & Zaheer, 2007; Muneepeerakul et al., 2013). Regional differences might result in some activities being closely related in some places but not in others, yet investigations of how cognitive proximity functions at the local level are still scarce (Boschma, 2017). A more context- and location-specific approach would avoid the bias that is created when proximity is considered as a fixed measure taken from databases at national or international levels, as is common in the existing proximity and relatedness literature.

Second, previous studies focus on developed countries, and little is known about how cognitive proximity and network structure influence innovation in developing countries (Whittle & Kogler, 2020). Developing countries generally have less-diversified productive structures (Imbs & Wacziarg, 2003), and relatively weak educational and institutional support frameworks (Hanushek, 2013) and physical and technological infrastructure (Eder & Trippl, 2019). This may result in low levels of local knowledge exchange, which, in turn, may hinder the innovation process (Grillitsch & Nilsson, 2015).

Third, especially in developing countries, micro and small enterprises (MSEs) are critical for the diffusion of innovations because they rely more on localised networks and informal communication (Acs & Audretsch, 1988; Kraemer-Mbula et al., 2019), taking advantage of the tacit knowledge exchanges that take place locally (Maskell & Malmberg, 1999; McCann, 2007). Although these firms often face financial and resource constraints to effective innovation (Nooteboom, 1994; Rogers, 2004), their greater dependence on local knowledge flows may positively influence the patterns of their cognitive proximity and their innovation performance.

In order to contribute to filling these research gaps, this study investigates whether cognitive proximity and network structure (measured as both degree centrality and betweenness centrality) matter for innovation in manufacturing MSEs in Johannesburg, South Africa. Our study contributes to the literature in four main ways. First, to compute their degree of cognitive proximity we consider three relevant dimensions that can bring firms closer. We develop and apply a new variable using the information on: (1) different skills that MSEs most need in their activities; (2) their technological relatedness represented by the sectors in which they operate; and (3) the types of external co-operation in which they engage. Our approach departs from previous studies that computed cognitive proximity between firms using only industrial classification. Our approach thus incorporates recent insights from the skill-relatedness approach (Content & Frenken, 2016; Whittle & Kogler, 2020) and the exchange of different types of knowledge, such as managerial, institutional, industry-specific and market knowledge (Alberti et al., 2021). Together, these three relevant dimensions result in a more nuanced, comprehensive and multidimensional measurement of cognitive proximity between firms.

The second key way in which we contribute to the literature is through our focus on a specific geographical area in a developing country to assess whether knowledge circulates among MSEs that are cognitively closer, and whether this influences innovation. We use a rich novel dataset focused on innovation and covering 711 manufacturing MSEs in Johannesburg. This helps to address the paucity of literature on this topic in Africa, as well as in developing countries more broadly. The context-specificity of networks and cognitive proximity make empirical evidence from different contexts especially important.

Thirdly, our study includes micro (not only small) enterprises, and informal (as well as formal) enterprises. The related empirical literature generally does not cover micro firms and informal firms, as these categories are excluded from most firm-level surveys. Both of these types of firms are especially important in developing country contexts, and furthermore are likely to have distinctive characteristics and patterns in terms of cognitive proximity, networks and innovation.

In our fourth contribution, we extend the literature on how cognitive proximity and network embeddedness affect innovation outcomes to consider the effects on the novelty of innovation. We thus separately analyse innovation outcomes in terms of whether the innovation is new to the firm, market, community or world. This sheds light on how cognitive proximity and networks differentially affect innovations that are associated more with the diffusion of previous innovations, from new scientific and technological breakthroughs. Taking into account the degree of innovation novelty brings nuance and deeper insights to our understanding of the relationship between firms' cognitive proximity and network position, and their innovation performance.

The results suggest that the degree of cognitive proximity between MSEs and their network position increase the likelihood of innovations being introduced, especially those related to the diffusion process that is associated more with MSEs. We find an inverted-U relationship between cognitive proximity and innovation outcomes, associated with the 'proximity paradox', in which proximity enhances the probability of innovating but too much proximity may lead to a cognitive lock-in. In terms of firms' network positions, we observe interesting differences by dimension (degree and betweenness centrality), by type of innovation (overall, product, and process/service), and by degree of innovation novelty.

The article proceeds as follows. Section 2 provides an overview of the relevant theoretical framework and literature. Section 3 discusses the methodology and novel data used in the analysis. The results are presented and discussed in Section 4, and Section 5 concludes.

2. Theoretical framework and literature

2.1 Cognitive proximity and innovation

Innovation can be understood as a process in which social actors create, absorb and recombine different sets of knowledge to generate new products and processes, which may be proximate to the extant knowledge (Nelson & Winter, 1982; Weitzman, 1998). This process depends on the coordination capacity of firms in selecting workers with related bases of knowledge and their ability to recombine them. As not all types of knowledge may readily be recombined, the concept of proximity is useful for understanding how agents deal with problems of coordination and uncertainty, and increase their capacity for learning from others (Boschma, 2005; Torre & Rallet, 2005). Proximity refers to the closeness of social actors with similar characteristics (Gertler, 2003) or the same system of representations (Torre & Rallet, 2005), which facilitates information transmission and knowledge exchanges (Bathelt et al., 2004; Storper & Venables, 2004). Cognitive proximity is associated with firms having similar levels of knowledge and expertise, which allows for the effective communication, learning, absorption, creation and diffusion of knowledge (Boschma, 2005; Nooteboom, 2000). There is also evidence suggesting that cognitive proximity is often related to increased levels of innovation (Fitjar et al., 2016; Nooteboom et al., 2007).

Cognitive proximity between firms is dynamic and is modified through continuous social interactions, which improve their ability to better understand the local sets of information (Nooteboom, 2000). Hence, repeated meetings and other interactions (whether organised or unplanned) among firms' managers and workers can diffuse ways of solving specific problems or share tacit knowledge (Bathelt et al., 2004; Storper & Venables, 2004). These interactions will shorten the cognitive distance between firms and increase both their learning capacity and the stock of knowledge that can be exchanged socially.

It is important to note that, although cognitive proximity facilitates effective communication and knowledge absorption, too much proximity might not improve the innovation capacity of firms, and in fact may even harm it. 'Excessive' knowledge similarity between firms may undermine the novelty of the information they exchange and lead to a 'cognitive lock-in', in the sense that they might not be able to recognise new opportunities to innovate. In sum, cognitive proximity between firms enhances their knowledge exchange, but the innovation process requires a certain level of dissimilarity between the types of knowledge to be recombined. Boschma and Frenken (2010) refer to this as the 'proximity paradox', and its influence on innovation capacity will depend on the optimal level of cognitive proximity between firms. We thus expect cognitive proximity between firms to foster learning and knowledge transmission, but that excessive cognitive proximity can diminish this (in particular by reducing the degree of novelty in the information that the firms exchange). This suggests that the relationship between innovation and cognitive proximity follows an inverse U-shaped pattern.

Based on the above discussion, we hypothesise that:

Hypothesis 1: Firms close to other firms in terms of cognitive proximity are more likely to introduce innovations, whereas too much cognitive proximity is negatively associated with the probability of innovating.

2.2 The knowledge network structure and innovation

An economy can be represented as an intricate web in which cognitive proximity is a driving force that brings firms with similar knowledge bases together and facilitates the transmission of complex information and tacit knowledge. Links between cognitively related firms enable mutual learning, coordinate specialised knowledge flows, and promote innovation (Hidalgo et al., 2007; Shutters et al., 2016).

However, access to relevant knowledge flows is uneven. Geographical proximity can provide opportunities for otherwise unconnected firms and workers to get to know each other, but not all the knowledge is uniformly available for all firms in the same location (Fitjar & Rodríguez-Pose, 2017). This differentiated access to relevant knowledge flows results from social interactions that shorten cognitive distances between some firms in the network and increase them for others. From this perspective, the latter types of firms are in the peripheral regions of the network with limited access to knowledge, while the former types of firms are located at the densest areas where local 'buzz' takes place (Gertler, 2003; Storper & Venables, 2004). Through these interactions, firms occupying central positions in the local network have more opportunities to access the knowledge required to innovate and increase their propensity for introducing innovations (Gilsing et al., 2008; Zaheer & Bell, 2005).

The degree of a firm's centrality in a network is another important factor in our framework. Degree centrality can be defined as the number of links a firm has with other firms. A firm establishes links with other, cognitively related firms, and the length of these links varies according to the cognitive distance between them. Establishing a new tie provides a firm with an opportunity to access additional sets of knowledge that can be useful for its innovation processes. In addition, a firm embedded at the densest areas of the network paths is located between other sets of firms. The higher this 'betweenness centrality' of a firm is, the greater the benefits from knowledge spillovers and the greater the quantity of knowledge it can access (Stuck et al., 2016).

This leads to our next hypothesis:

Hypothesis 2: The more central the position of a firm in a knowledge network (degree centrality and betweenness centrality), the higher its probability of introducing innovations.

Therefore, to better understand the process of innovation, we have to consider the cognitive proximity that brings firms closer and positions them in central areas of the network where related knowledge flows and which enhances their capacity for introducing new products, processes and services.

3. Methodology and data

3.1 Model estimation

In order to assess how cognitive proximity between firms and their position in the local knowledge networks affect the probability of introducing innovations, we estimated binomial logistic regressions in which innovation was the dependent variable. Firms were asked if they had introduced entirely new or significantly improved products, processes and/or services.

Innovation is binary and takes a value of 1 if a firm has introduced any of the three innovation types, and zero otherwise.

The logistic model can be expressed as

$$P(Y) = \frac{1}{1+e^{Z}} P, \qquad (1)$$

where Y is the dependent variable and Z is the linear combination of independent and control variables. Thus, the model takes the following form:

$$P(Innovation = 1) = \alpha_i + \beta_1 \cdot Cognitive\ Proximity_i + \beta_2 \cdot Cognitive\ Proximity_i^{\,2} + \beta_3 \cdot Degree\ Centrality_i + \beta_4 \cdot Betweenness\ Centrality_i + \beta_i \cdot Controls + \varepsilon_i, \quad (2)$$

As errors for firms belonging to the same industry may be correlated, the regression results were adjusted using heteroscedasticity-robust standard errors clustered at the industry level (Cameron et al., 2011; Wooldridge, 2003).

Based on our hypotheses, we expect *cognitive proximity*, *degree centrality* and *betweenness centrality* to show positive and statistically significant coefficients, and that *cognitive proximity^2* would show a negative and statistically significant coefficient for the likelihood of introducing innovations. Section 3.3 sets out how we measured cognitive proximity; section 3.4 explains both our measures of firms' network position (degree centrality and betweenness centrality); while section 3.5 discusses our control variables. Below, section 3.2 provides an overview of our dataset.

3.2 Data

Innovation in MSEs derives in part from the fact that MSEs are well adapted to local 'buzz' and have a more flexible organisational structure to benefit from accessing tacit and sticky knowledge that circulates locally. Indeed, Kraemer-Mbula et al. (2019) have shown that informal communication and knowledge exchange are crucial for developing innovation capacity in African countries. For these reasons, Johannesburg offers an ideal context for providing evidence on local knowledge flows, interactive learning and innovation: its economic relevance promotes agglomeration economies through localisation externalities that attract workers with higher levels of education, increase local productivity and intensify the flows of goods and knowledge (Krugell & Rankin, 2012). The City of Johannesburg is the largest and most economically developed city in the country, accounting for 10% of South Africa's population and 16% of GDP (City of Johannesburg, 2021).

We use data from a rich new survey of manufacturing MSEs in Johannesburg.[1] The survey focused on innovation, and covered both formal and informal enterprises. This survey is the most comprehensive ever conducted in South Africa for understanding the extent and impact of firm-level innovation, with few such surveys in other developing countries, especially in Africa. Unlike most other surveys, especially other innovation surveys, this survey includes micro enterprises as well as informal enterprises.[2] Given the importance of both of these types of enterprises in developing countries, especially in Africa, this dataset enables a valuable extension of the empirical literature on networks and cognitive proximity to cover all MSEs

[1] The survey was undertaken under the DSI/NRF South African Research Chair in Industrial Development, through the project 'Community of Practice in Innovation and Inclusive Industrialisation'.

[2] For example, the World Bank Enterprise Surveys, which are the most widely used sources of firm-level data internationally, especially in developing countries, exclude micro-enterprises (with fewer than five employees) and informal enterprises.

(including the smallest firms with fewer than five employees), and both formal and informal enterprises.

The survey comprised 74 questions, covering a range of aspects of firm characteristics and behaviour, with a focus on innovation (see Table A.1 for the questions utilised in this study). Firms were asked about their demographic and business profiles, innovation processes, training and skills development, linkages, and financial dimensions. The survey questions on innovation followed the Oslo Manual (OECD/Eurostat, 2018).

The survey covered all enumerator areas (EAs) classified for commercial or industrial use in the urban development zone (UDZ) of the City of Johannesburg, any additional EAs classified as industrial in the most recent (2011) Census, and contiguous residential EAs where manufacturing activities had extended across EA boundaries. The survey thus covered 200 EAs, which included all 142 industrial EAs in the City of Johannesburg and all 58 commercial EAs in the UDZ.

In all these EAs, the first step in the data collection was a census of all businesses, from which non-manufacturing enterprises and those with more than 50 employees were excluded. Interviews were conducted face to face with 724 firms by trained fieldworkers from a professional survey company, over the period June to August 2019. After the data-cleaning process, 711 interviews were deemed to be usable. The data is representative of micro and small manufacturing firms in Johannesburg.

3.3 Measuring cognitive proximity

To measure cognitive proximity between MSEs, we introduced a new variable based on three different dimensions. First, we used the industrial classification to compute technological relatedness, as is common in the related variety literature (Frenken et al., 2007). MSEs operating in the same industry share technological expertise that is associated with common organisational knowledge and similar ways of solving operational problems and day-to-day issues. Technologically related firms are often better prepared for benefiting from local knowledge spillovers (Boschma & Frenken, 2012), which enhances their innovation output (Castaldi et al., 2015; Frenken et al., 2007). All firms in the dataset were classified within 25 manufacturing sectors, providing information about their technological relatedness.

Second, we also considered the types of skills most needed in the business, in line with the most recent skill-relatedness approach (Content & Frenken, 2016; Whittle & Kogler, 2020). Knowledge exchange between MSEs is facilitated if they employ workers sharing similar skills. Some studies show that firms' proximity in the skills space has influenced regional diversification and development (Fitjar & Timmermans, 2017; Neffke & Henning, 2013). Firms were asked to indicate the three most needed skills in their business out of a list with 11 different skills. Thus, based on the responses, we were able to link each firm to others in terms of skill similarity.

Third, we included the extent of co-operation with other MSEs in the same industry based on different types of knowledge. MSEs are closer in a cognitive dimension if they engage in external co-operation that involves exchanging similar types of knowledge. Interactions aimed at sharing information of general interest strengthen the cognitive proximity and enhance the learning processes between MSEs (Bathelt et al., 2004; Maskell & Malmberg, 1999). Based on the responses, we were able to classify MSEs that exchange similar types of knowledge as similar.

Together, these three dimensions result in a more nuanced, comprehensive and multidimensional measurement of cognitive proximity between firms than in the extant literature.

Having identified the three dimensions to be included in the new variable measuring cognitive proximity, the next step was to arrange this information in a matrix (43 x 711). To obtain the cognitive proximity, φ, between firms, we followed Hidalgo et al. (2007) and computed the minimum of the pairwise conditional probability of firms i and j sharing a set of knowledge based on skills, technological relatedness and the type of co-operation. We calculated the number of normalised co-occurrences between firms from the matrix and compared them to the random probability of co-occurrence, according to:

$$\varphi_{i,j} = \min\{P(knowledge_i)|P(knowledge_j), P(knowledge_j)|P(knowledge_i)\}. \quad (3)$$

Two firms are closer or related in the cognitive dimension if $\varphi > 1$, and are distant or unrelated if $\varphi < 1$. We analysed pairs of cognitively close MSEs ($\varphi > 1$) in order to focus on the relevant areas of the network in which the likelihood of knowledge spillovers is greater. Thereafter, we averaged φ values across MSEs and rescaled cognitive proximity between 1 to 100.

3.4 Measuring the network position of firms

Based on our measure of cognitive proximity, Figure 1 shows the knowledge network connecting MSEs in Johannesburg, clustered by industry classification. Each node in the network graph represents a manufacturing sector, with shaded nodes indicating innovative MSEs, while white nodes reflect non-innovative firms. The node size reflects the number of connections that firms in an industry have with others. The lines connecting nodes are cognitive linkages.

The configuration of industries shows that, in general, more cognitively related industries are located closer together in the network. For example, innovative MSEs in the ‘motor vehicles’ industry are more closely related to innovative MSEs in ‘electronics’ than to innovative firms in the ‘textiles’ industry. In our framework, this implies that knowledge spillovers should be stronger between firms in the ‘motor vehicles’ and ‘electronics’ sectors than between firms in one of these industries and in the ‘textile’ industry. We also observe that, with a few exceptions, innovative MSEs tend to be closer to other innovative ones and further from non-innovative MSEs.

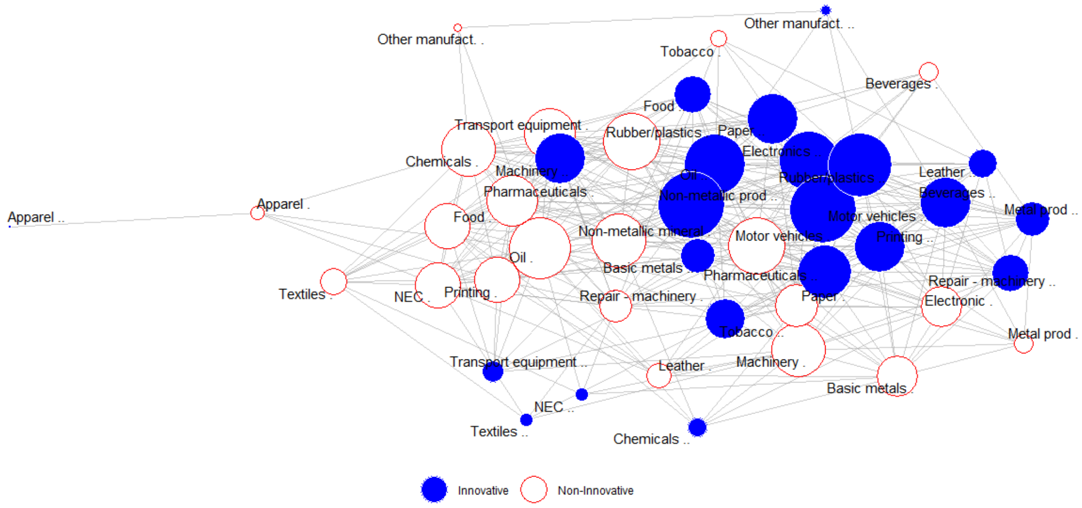

Figure 1: Knowledge networks among sample firms

Source: Authors' elaboration using survey data.

In order to identify an MSE's position in the network, we computed two variables: degree centrality and betweenness centrality. These are included in our regressions as possible determinants of innovation outcomes.

Degree centrality refers to the number of firms that an MSE is directly connected with. MSEs with high degree centrality are more involved with, and tend to have a great influence on, directly linked firms (Stuck et al., 2016). This variable is simply the number of edges incident to it or the number of its connections. Thus, to compute the degree d_v of a node v in a network graph $G = (V, E)$, all that is required is for the number of edges (links) in E incident upon v to be counted (Coscia, 2021).

The second variable, betweenness centrality, refers to the extent to which an MSE is located between other MSEs. It is calculated as the fraction of the shortest paths between other firms that pass through the local MSE (Stuck et al., 2016). To compute the betweenness centrality (C_b) of a firm, we counted the number of paths passing through node v, where v is neither an origin nor a destination, according to:

$$C_b = \sum_{s \neq t \neq v \in V} \frac{\sigma(s, t|v)}{\sigma(s,t)}, \quad (4)$$

where $\sigma(s, t|v)$ is the total number of shortest paths between s and t passing through v, and $\sigma(s, t)$ is the total number of shortest paths between s and t, independent of whether or not they pass through v (Coscia, 2021). Both network variables were rescaled to between 1 and 100.

3.5 Control variables

Successful innovation outcomes also depend on other determinants. The first set of variables we added as controls are related to the MSEs' internal resources and absorption capacity (Cohen & Levinthal, 1990). For example, the entrepreneur's educational level is associated with a higher rate of innovation and adoption of new technology (Koellinger, 2008; Martínez-Román & Romero, 2017). We included the education level of managers on a scale ranging from 1 (no formal schooling) to 10 (postgraduate degree). We also controlled for the years of work experience that managers have in the sector and the age of the business. Since MSEs' innovation capacity is expected to be positively associated with R&D activities (Baumann & Kritikos, 2016; Martínez-Román & Romero, 2017), we added a binary variable related to R&D that takes the value of 1 if the firm engaged in R&D activities, and zero otherwise.

We also controlled for different dimensions of training and skills, which are key elements in innovation capacity. Indeed, skilled labour seems to be more important to innovation among micro-firms relative to other firms (Andersson & Lööf, 2012). Innovation requires a range of different skills, which are typically more scarce for MSEs (Rogers, 2004). Thus, finding workers with the required levels of creative and innovative skills can be difficult (Faherty & Stephens, 2016), especially for activities in which the skills needed to innovate change rapidly. We added a variable related to the extent to which the MSE faces a skills constraint. Skills availability provides information that ranges from level 1, when it is *very difficult* to find workers with the required skills, level 2, when it is *fairly difficult* to find them, level 3, when it is not *very difficult* to find them, to level 4, when it is *not at all difficult* to find employees with the required skills. We also controlled for the capacity of MSEs to retain qualified workers. Workers usually learn and become more skilled when performing their daily tasks, and other firms can poach them from the original workplace. As a result, failure to retain workers may affect MSEs' innovative performance (Gertler, 2003; Nooteboom et al., 2007). We added a variable related to the ease with which a firm retains employees with the required skills, ranging between the same levels presented earlier – from 1 (*very difficult*) to 4 (*not at all difficult*). Regarding the degree to which skills needed from workers change over time, we added a variable related to skill stability. This variable ranges from level 1, when skills *change very quickly*, to level 2, when skills *change fairly quickly*, level 3, when skills *do not change very quickly*, and level 4, when skills *do not change at all*. Finally, investments in human capital are necessary to improve the workers' skills and, for this reason, providing training activities for employees is associated with higher innovation capabilities (Martínez-Román & Romero, 2017; Rogers, 2004). We added a binary variable that takes a value equal to 1 if a firm's workers participate in training, learn from more experienced colleagues and supervisors, or participate in structured apprenticeships, and zero if workers learn by doing through trial and error or by self-driven learning through the internet.

The second set of variables we added as controls is related to the establishment of external connections with different actors located within and outside the region. MSEs typically require regular external connections to compensate for the lack of internal expertise and limited resources (De Jong & Vermeulen, 2006; Nooteboom, 1994). Although these firms generally have only limited collaboration with universities and research centres (Martínez-Román & Romero, 2017), there is evidence that spillovers from the research activities of universities have a positive effect on their innovative performance (Acs et al., 1994). In addition, MSEs are efficient at interacting closely with partners in an informal way and collecting information from their daily contacts with different customers, suppliers, trade associations and government authorities (Andersson & Lööf, 2012; Gronum et al., 2012; Moilanen et al., 2014; Nooteboom, 1994).

We controlled for the type of customers from which a firm gets most of its sales by adding dummy variables for individuals (baseline), business, and government/public institutions. We also controlled for customer location in terms of whether most of a firm's customers are located in South Africa (baseline) or abroad. We further added information about the number of suppliers. Firms indicated the number of suppliers, ranging from level 1 (none) to level 6 (more than 100). We controlled for firms' participation in business associations and inter-firm co-operation. In addition, we included dummy variables for 25 manufacturing sectors to control for heterogeneity across industries.

4. Results

4.1 Descriptive statistics

Descriptive statistics are shown in Table 1. (The names, types and definitions of all the variables are set out in Table A.1 in the Appendix.)

We hypothesised that MSEs in Johannesburg will benefit from their cognitive proximity and position in local knowledge networks to innovate. The descriptive statistics suggest that innovative MSEs present a high level of cognitive proximity, indicating that they are located closer to other MSEs with similar cognitive bases and are expected to have better opportunities to exchange knowledge, learn and innovate. Innovative MSEs are often located at central positions in the network, as they present higher levels of degree centrality and betweenness centrality than do non-innovative firms. These figures suggest that innovative MSEs potentially could reach higher levels of knowledge that can be used for novel combinations.

These issues are analysed in the econometric results that follow. Our main results are discussed in section 4.2, while section 4.3 extends this analysis with the dimension of innovation novelty.

Table 1: Descriptive statistics

	All MSEs					Innovative MSEs		Non-innovative MSEs	
	N	Min	Max	%		N	%	N	%
Innovation	711	0	1	100		382	53.73	329	46.27
Product innovation	711	0	1	100		347	48.80	364	51.20
Process and service innovation	711	0	1	100		104	14.63	607	85.37
New to the business	711	0	1	100		154	21.66	557	78.34
New to the community	711	0	1	100		92	12.94	619	87.06
New to the market	711	0	1	100		111	15.61	600	84.39
New to the world	711	0	1	100		25	3.52	686	96.48
	All MSEs					**Innovative MSEs**		**Non-innovative MSEs**	
	N	Min	Max	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Cognitive proximity	711	1	100	21.66	12.18	21.77	11.56	21.52	12.88
Cognitive proximity^2	711	1	10,000	617.2	803.78	607.38	699.36	628.54	911.11
Degree centrality	711	1	100	43.05	9.7	43.12	9.65	42.97	9.77
Betweenness centrality	711	1	100	5.08	5.88	5.64	6.95	4.42	4.22
Education	711	1	10	5.78	2.01	6.34	1.9	5.13	1.95
R&D	711	0	1	0.15	0.35	0.21	0.41	0.07	0.26
Skills – availability	711	1	4	2.04	1.14	1.9	1.07	2.2	1.2
Skills – retaining	711	1	4	2.37	1.11	2.34	1.1	2.41	1.12
Skills – training	711	0	1	0.79	0.41	0.81	0.39	0.75	0.43
Skills – stability	711	1	4	2.21	1.09	2.09	1.03	2.34	1.14
Experience	711	1	53	13.27	10.56	13	10.67	13.6	10.44
Age of business	711	2	119	14.85	13.99	14.16	12.84	15.64	15.2
Customer – individuals	711	0	1	0.73	0.44	0.71	0.46	0.77	0.42
Customer – business	711	0	1	0.26	0.44	0.28	0.45	0.23	0.42
Customer – government	711	0	1	0.01	0.08	0.02	0.12	0.00	0.00
Customer – local	711	0	1	0.95	0.21	0.93	0.25	0.97	0.16
Customer – abroad	711	0	1	0.05	0.18	0.07	0.22	0.03	0.12
# Suppliers	711	1	6	2.07	0.83	2.2	0.83	1.92	0.8
Member – business association	711	0	1	0.09	0.29	0.1	0.3	0.09	0.28
Co-operation	711	0	1	0.48	0.5	0.55	0.5	0.39	0.49

4.2 Main results

Table 2 reports our econometric results for overall innovation (column 1), product innovation (column 2), and process and service innovation (column 3). In each case, we are interested in the effects of cognitive proximity and of firms' network positions (degree centrality and betweenness centrality) on these innovation outcomes.

Regarding the association between innovation and MSEs' position in a network, we find mixed results. Degree centrality – the number of connections that an MSE has in the network – has the expected result only for product innovation. For this type of innovation, which was more common in our sample, connecting with other firms allows access to new sets of knowledge that enhance innovation. However, degree centrality was not statistically significant for overall innovation and presented an unexpected negative signal for process and service innovation. A possible explanation for the latter result could be that, in our sample of firms, those with more connections in the network become more accustomed to doing things in the same way and less innovative in their production processes; this could be explored further in future research.

When looking at the measure of betweenness centrality, we found more consistent results for overall and product innovation. Both estimations show a positive and statistically significant coefficient on betweenness centrality, indicating that being more centrally positioned at the middle of a network, where knowledge circulates, increases the likelihood of introducing innovations. However, we did not find any influence of betweenness centrality on process and service innovation. We thus found full evidence for our *Hypothesis 2* only for product innovation, with partial evidence for overall innovation, where only betweenness centrality seems to play a significant role.

The results from our first set of controls, related to MSEs' internal resources, are mostly as expected. The likelihood of introducing all types of innovation increases for MSEs that hire more educated managers and engage in R&D activities. The negative and significant coefficients for skills availability indicate that firms that struggle to find employees with the required skills also have a higher probability of innovating. This is likely because such firms have higher skills needs, requiring skills that are not readily available in the local labour market. We find a positive influence of training on overall and product innovation. However, for these innovation types, we find that the probability of innovating increases for MSEs that hire workers with skills that change quickly.

Table 2: Logistic regressions: overall innovation, product innovation, and process and service innovation

	Overall innovation (1)	Product innovation (2)	Process and service innovation (3)
Cognitive proximity	0.065*** (0.017)	0.051*** (0.014)	0.078*** (0.029)
Cognitive proximity²	-0.001*** (0.000)	-0.001** (0.000)	-0.002** (0.001)
Degree centrality	0.008 (0.008)	0.016* (0.009)	-0.040* (0.022)
Betweenness centrality	0.056** (0.025)	0.056** (0.023)	0.010 (0.016)
Education	0.320*** (0.057)	0.264*** (0.063)	0.350*** (0.047)
R&D	0.820*** (0.290)	0.659* (0.338)	0.980*** (0.307)
Skills – availability	-0.326*** (0.096)	-0.244** (0.108)	-0.272** (0.113)
Skills – retaining	0.115 (0.133)	0.143 (0.100)	0.044 (0.142)
Skills – training	0.337** (0.167)	0.329* (0.192)	0.224 (0.239)
Skills – stability	-0.201** (0.093)	-0.213** (0.087)	-0.085 (0.115)
Experience	0.003 (0.007)	-0.004 (0.010)	0.024 (0.018)
Age of business	-0.007 (0.007)	-0.006 (0.005)	-0.010 (0.011)
Customer – business	0.086 (0.225)	-0.071 (0.242)	0.363 (0.298)
Customer – government	14.308*** (0.598)	-0.396 (0.888)	1.655 (1.222)
Customer – abroad	1.410*** (0.523)	1.310*** (0.332)	-1.669* (0.928)
# Suppliers	0.255** (0.106)	0.217* (0.116)	0.192 (0.141)
Member – business association	-0.373* (0.193)	-0.391** (0.187)	-0.277 (0.253)
Co-operation	0.402** (0.196)	0.334 (0.212)	0.028 (0.292)
Constant	-3.529*** (0.687)	-3.539*** (0.686)	-3.364*** (1.230)
Observations	711	711	711
Industry dummies	Yes	Yes	Yes
Nagelkerke R^2	0.294	0.239	0.252
Chi-square	176.42***	140.33***	109.05***

Note: Heteroskedasticity-robust standard errors are shown in parentheses. Coefficients are statistically significant at * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

For the external relationship dimension, the probability of introducing overall innovation is higher for MSEs that sell mainly to government and public institutions, compared to firms that sell primarily to individuals and businesses. This result highlights the relevant role that public procurement plays in innovation, especially in developing countries (Rocha, 2018). MSEs with most of their customers located abroad have a higher likelihood of introducing overall and product innovation. However, we find that MSEs that innovate in processes and services have most of their customers in South Africa. This could suggest that selling to customers abroad requires more innovative capacity to adjust products to follow regulations and patterns in international markets. In contrast, more connections with local customers seem to be associated with the diffusion of non-product innovations. For domestic markets, process innovations that improve the efficiency of production may be more important than developing new products. The number of suppliers also increases the probability of introducing overall and product innovation, while co-operation with firms in the same sector seems to increase the probability of introducing only overall innovation. Contrary to our expectations, being a member of business associations seems to negatively affect overall and product innovation. This may be because MSEs are cautious about revealing their innovation plans in a context where there are many competitors, such as business associations. The more dynamic and innovative firms that aim to do things differently possibly prefer to operate independently of established business associations.

4.3 Extension: novelty of innovation

Next, we extend the analysis to investigate if the relationship between cognitive proximity and the network position of MSEs and their innovation outcomes varies for different degrees of innovation novelty. This extension is motivated in part by a recognition of the different roles typically played by small and large firms in innovation systems, with our focus here being on MSEs. Large firms generally tend to play a major role in scientific and technological breakthroughs by employing specialised teams focused on science and R&D-based innovations. In contrast, small firms commonly play a particular role in more incremental improvements in existing technologies, their implementation, application, differentiation and adaptation, thus in the diffusion of innovations (Nooteboom, 1994). Furthermore, consideration of the degree of novelty brings nuance and deeper insights to our understanding of the relationship between firms' cognitive proximity and network position and their innovation performance.

The diffusion of innovations is captured by both 'new to the business' and 'new to the community' innovations. In contrast, innovations that are 'new to the market' refer to those implemented for the first time in the market or sector, and those 'new to the world' represent the introduction of innovation for all markets and industries, domestic and international (OECD/Eurostat, 2018). Our survey data indicates that 40% of the innovations are 'new to the business', 24% are 'new to the community', 29% are 'new to the market', and only 7% are 'new to the world'. Considering different degrees of novelty, we expect that cognitive proximity between MSEs and their embeddedness in the knowledge network might be associated more with the diffusion of innovation rather than innovations with a greater degree of novelty.

Table 3 shows the estimates for the extended model, following an increasing order of novelty in columns 1 to 4. The inverted-U relationship between cognitive proximity and innovation (as per hypothesis 1) is evident only for ‘new to the business’ and ‘new to the community’ innovations. When we consider more radical innovation that is ‘new to the market’, we did not find a role for cognitive proximity, and this variable is only positive and significant for its quadratic term for ‘new to the world’ innovations. A possible explanation for these results might be that the introduction of innovations with greater degrees of novelty relies more on the combination of unrelated knowledge sets than related ones that circulate locally (Castaldi et al., 2015), including possibly through cognitive proximity to larger firms and firms outside of the local area.

In terms of the network position of MSEs, we find a positive and significant influence of betweenness centrality on ‘new to community’ innovations and of degree centrality on ‘new to the world’ innovations. However, this latter variable shows a negative association with the introduction of ‘new to the market’ innovations. Looking at the MSEs’ network embeddedness and connections with customers abroad, we can suggest that firms that introduced ‘new to the community’ innovations benefited more from local ties and knowledge from their domestic customers than from contacts abroad. Conversely, MSEs that introduced ‘new to the market’ innovations seemed to benefit more from connections with customers located abroad than from the local network. However, the results for ‘new to the world’ innovations do not allow us to confirm this hypothesis: (local) degree centrality seems to be more relevant for the likelihood of introducing innovations. These differences in the results could be derived from the relatively few events related to ‘new to the world’ innovation and require further research. Despite the inconclusive results for innovation with higher degrees of novelty, we have confirmed that cognitive proximity and betweenness centrality are important factors for enabling more efficient diffusion of innovations.

Table 3: Logistic regressions: Degree of innovation novelty

	New to the business (1)	New to the community (2)	New to the market (3)	New to the world (4)
Cognitive proximity	0.080^{***} (0.023)	0.132^{***} (0.048)	-0.039 (0.041)	-0.040 (0.049)
Cognitive proximity^2	-0.001^{**} (0.000)	-0.002^{***} (0.001)	0.000 (0.001)	0.001^{**} (0.001)
Degree centrality	0.004 (0.012)	0.001 (0.020)	-0.022^{***} (0.007)	0.040^{*} (0.024)
Betweenness centrality	0.016 (0.021)	0.048^{**} (0.019)	-0.004 (0.018)	0.024 (0.016)
Education	0.273^{***} (0.043)	-0.032 (0.044)	0.141^{**} (0.069)	0.364^{*} (0.213)
R&D	0.078 (0.249)	0.422 (0.429)	0.367 (0.273)	1.044^{**} (0.502)
Skills – availability	-0.161^{**} (0.064)	-0.267^{**} (0.116)	-0.078 (0.138)	-0.567^{***} (0.143)
Skills – retaining	0.210 (0.145)	-0.245^{**} (0.106)	0.059 (0.089)	0.190 (0.139)
Skills – training	0.586^{*} (0.324)	-0.023 (0.340)	-0.212 (0.220)	0.126 (0.355)
Skills – stability	-0.011 (0.064)	-0.361^{***} (0.098)	0.063 (0.141)	-0.396^{***} (0.095)
Experience	-0.010 (0.014)	-0.002 (0.014)	0.001 (0.010)	0.038^{**} (0.018)
Age of business	0.007 (0.007)	-0.012 (0.015)	-0.007 (0.009)	-0.064^{**} (0.031)
Customer – business	0.226 (0.195)	-0.300 (0.340)	0.153 (0.186)	0.295 (0.253)
Customer – government	-0.428 (1.421)	0.063 (0.571)	1.111 (1.122)	-15.239^{***} (1.421)
Customer – abroad	0.617 (0.813)	-1.485^{**} (0.649)	0.990^{*} (0.600)	1.265 (1.378)
# Suppliers	0.114 (0.119)	0.319 (0.213)	0.098 (0.079)	-0.453^{**} (0.223)
Member – business association	-0.444 (0.515)	-0.114 (0.536)	-0.443 (0.539)	1.382 (1.189)
Co-operation	0.317 (0.254)	0.306 (0.263)	0.011 (0.248)	-0.049 (0.508)
Constant	-5.097^{***} (0.836)	-1.347 (1.205)	-2.879^{***} (0.701)	-23.729^{***} (1.545)
Observations	711	711	711	711
Industry dummies	Yes	Yes	Yes	Yes
Nagelkerke R^2	0.183	0.209	0.128	0.392
Chi-square	89.734^{***}	84.753^{***}	54.909^{*}	77.32^{***}

Note: Heteroskedasticity-robust standard errors are shown in parentheses. Coefficients are statistically significant at * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

5. Conclusion

This study investigates whether cognitive proximity and knowledge network structure matter for MSEs' innovation. We used new data from a survey of 711 manufacturing MSEs located in Johannesburg, South Africa. Drawing on theoretical insights from the proximity literature (Boschma, 2005; Boschma & Frenken, 2010; Torre & Rallet, 2005), we introduced a multidimensional measure of cognitive proximity between firms and applied it to map a knowledge network in which we could position each one. We tested the association of cognitive proximity and network embeddedness with the introduction of innovations.

One way in which this article contributes to the literature is thus through the development and application of a new, three-dimensional measure of cognitive proximity. We computed our main variable by using information on different skills that MSEs need most in their activities, their technological relatedness represented by the sectors in which they operate, and the types of external co-operation in which they engage.

Our application of this variable further contributes to the literature by providing empirical evidence for the influence of cognitive proximity and localised knowledge network structures on innovation outcomes in a developing country, where existing empirical evidence is sparse – especially in the African context. Our focus on MSEs is especially germane, given the lack of evidence for these firms, in particular micro-firms and informal firms that are typically excluded from firm-level surveys and hence from the related literature, yet are very important in developing countries.

In addition to testing how cognitive proximity and knowledge network centrality affect overall innovation, product innovation and process innovation, we also extend the literature by considering differential outcomes in terms of the novelty of innovation. We ran separate regressions in which the dependent variables were innovations that were new to the business, community, market and world.

Our findings suggest that the relationship between cognitive proximity and innovation follows an inverse U-shaped curve, associated with the 'proximity paradox'. Proximity allows for efficient knowledge exchange that enhances the probability of innovating, but too much proximity undermines the novelty of the knowledge exchanged. This may lead to a cognitive lock-in, thus potentially diminishing the MSEs' innovation capacity. We found evidence that more centrally located MSEs in the network are better positioned to access new knowledge sets and increase their likelihood of introducing innovation.

Our extended results reveal interesting differences in terms of the novelty of innovation. The inverted-U relationship found for cognitive proximity in our main results only holds for innovations that are new to the firm or community; that is, it is associated more with the diffusion of innovations. It appears that more radical innovations, which are new to the market or the world, depend less on locally circulating knowledge sets. In terms of firms' network position, we find degree centrality to be particularly important to innovations that are new to the world, and betweenness centrality to be important to innovations that are new to the community.

Our results may have some important policy implications. First, as cognitive proximity matters for innovation, policy intervention should focus on developing and supporting the local stock of related knowledge that can more easily be recombined. However, the proximity paradox suggests that the local knowledge space cannot be too narrow because of the risk of cognitive lock-in. Thus, policies oriented to increasing the stock of local knowledge also have to pay attention to ways of diversifying it. A greater variety of related knowledge at the regional level is associated with better economic performance (Frenken et al., 2007), the emergence of new industries (Boschma et al., 2013), and innovation (Castaldi et al., 2015). Second, policies oriented to strengthen the connections among firms' managers and workers, and between them and other social actors, should also be considered. Trade missions, sectoral industry fairs and workshops organised by government and business associations could result in more frequent interactions and greater access to novel information. Third, policies could encourage new financial mechanisms to increase the internal resources and the R&D capacity of MSEs. Finally, it is important that policies promote investments in physical infrastructure and digital technologies, which are essential for establishing external pipelines to universities, research institutions, suppliers and customers within and outside the region.

This study is not without its limitations. While our dataset is rich and novel, it provides a static snapshot, whereas the mechanisms underlying learning and innovations are dynamic and can change over time. We expect that further surveys can add a temporal dimension to increase our understanding of the dynamic evolution of local knowledge. Our analysis also focuses on an informal relational network (Hidalgo et al., 2007), rather than on assessing a formal network made especially for collaborative innovations (Gilsing et al., 2008). Finally, we focus on cognitive proximity, with geographical proximity already established because firms are within the geographical limits of Johannesburg. Additional studies might investigate the interactions between cognitive and spatial proximities, while new surveys can add new types of proximity, considering the role of organisational, institutional and social proximities in innovation processes.

References

Acs, Z. J., & Audretsch, D. B. (1988). Innovation in large and small firms: An empirical analysis. *The American Economic Review, 78*(4), 678–690.

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1994). R&D spillovers and recipient firm size. *The Review of Economics and Statistics, 76*(2), 336–340.

Alberti, F. G., Belfanti, F., & Giusti, J. D. (2021). Knowledge exchange and innovation in clusters: A dynamic social network analysis. *Industry and Innovation, 28*(7), 880–901.

Andersson, M., & Lööf, H. (2012). Small business innovation: Firm level evidence from Sweden. *The Journal of Technology Transfer, 37*(5), 732–754.

Balland, P.-A., Belso-Martínez, J. A., & Morrison, A. (2016). The dynamics of technical and business knowledge networks in industrial clusters: Embeddedness, status, or proximity? *Economic Geography, 92*(1), 35–60.

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography, 28*(1), 31–56.

Baumann, J., & Kritikos, A. S. (2016). The link between R&D, innovation and productivity: Are micro firms different? *Research Policy, 45*(6), 1263–1274.

Bell, G. G., & Zaheer, A. (2007). Geography, networks, and knowledge flow. *Organization Science, 18*(6), 955–972.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies, 39*(1), 61–74.

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies, 51*(3), 351–364.

Boschma, R., & Frenken, K. (2010). The spatial evolution of innovation networks: A proximity perspective. In R. Boschma & R. Martin (Eds.), *The handbook of evolutionary economic geography* (120-135). Edward Elgar Publishing.

Boschma, R., & Frenken, K. (2012). Technological relatedness and regional branching. In H. Bathelt, M. Feldman, & D. F. Kogler (Eds.), *Beyond territory: Dynamic geographies of innovation and knowledge creation* (64-81). Routledge.

Boschma, R., & Ter Wal, A. L. J. (2007). Knowledge networks and innovative performance in an industrial district: The case of a footwear district in the south of Italy. *Industry & Innovation, 14*(2), 177–199.

Boschma, R., Minondo, A., & Navarro, M. (2013). The emergence of new industries at the regional level in Spain: A proximity approach based on product relatedness. *Economic Geography, 89*(1), 29–51.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics, 29*(2), 238–249.

Capone, F., & Lazzeretti, L. (2018). The different roles of proximity in multiple informal network relationships: Evidence from the cluster of high technology applied to cultural goods in Tuscany. *Industry and Innovation, 25*(9), 897–917.

Castaldi, C., Frenken, K., & Los, B. (2015). Related variety, unrelated variety and technological breakthroughs: An analysis of US state-level patenting. *Regional Studies, 49*(5), 767–781.

City of Johannesburg (2021). City of Johannesburg Integrated Development Plan 2021-26. Available at: https://www.joburg.org.za/documents_/Documents/2021-2016%20Final%20IDP/2021-26%20FINAL%20IDP%2021May%202021.pdf

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128–152.

Content, J., & Frenken, K. (2016). Related variety and economic development: A literature review. *European Planning Studies, 24*(12), 2097–2112.

Coscia, M. (2021). *The atlas for the aspiring network scientist*. IT University of Copenhagen.

De Jong, J. P. J., & Vermeulen, P. A. M. (2006). Determinants of product innovation in small firms: A comparison across industries. *International Small Business Journal, 24*(6), 587–609.

Eder, J., & Trippl, M. (2019). Innovation in the periphery: Compensation and exploitation strategies. *Growth and Change, 50*(4), 1511–1531.

Faherty, U., & Stephens, S. (2016). Innovation in micro enterprises: Reality or fiction? *Journal of Small Business and Enterprise Development, 23*(2), 349–362.

Fitjar, R. D., & Rodríguez-Pose, A. (2017). Nothing is in the air. *Growth and Change, 48*(1), 22–39.

Fitjar, R. D., & Timmermans, B. (2017). Regional skill relatedness: Towards a new measure of regional related diversification. *European Planning Studies, 25*(3), 516–538.

Fitjar, R. D., Huber, F., & Rodríguez-Pose, A. (2016). Not too close, not too far: Testing the Goldilocks principle of 'optimal' distance in innovation networks. *Industry and Innovation, 23*(6), 465–487.

Frenken, K., Oort, F. V., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies, 41*(5), 685–697.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there). *Journal of Economic Geography, 3*(1), 75–99.

Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., & Van den Oord, A. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy, 37*(10), 1717–1731.

Grillitsch, M., & Nilsson, M. (2015). Innovation in peripheral regions: Do collaborations compensate for a lack of local knowledge spillovers? *The Annals of Regional Science, 54*(1), 299–321.

Gronum, S., Verreynne, M., & Kastelle, T. (2012). The role of networks in small and medium-sized enterprise innovation and firm performance. *Journal of Small Business Management, 50*(2), 257–282.

Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review, 37*, 204–212.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science, 317*(5837), 482–487.

Imbs, J., & Wacziarg, R. (2003). Stages of diversification. *American Economic Review, 93*(1), 63–86.

Koellinger, P. (2008). Why are some entrepreneurs more innovative than others? *Small Business Economics, 31*(1), 21.

Kraemer-Mbula, E., Lorenz, E., Takala-Greenish, L., Jegede, O. O., Garba, T., Mutambala, M., & Esemu, T. (2019). Are African micro- and small enterprises misunderstood? Unpacking the relationship between work organisation, capability development and innovation. *International Journal of Technological Learning, Innovation and Development, 11*(1), 1–30.

Krugell, W., & Rankin, N. (2012). Agglomeration and firm-level efficiency in South Africa. *Urban Forum, 23*(3), 299–318.

Martínez-Román, J. A., & Romero, I. (2017). Determinants of innovativeness in SMEs: Disentangling core innovation and technology adoption capabilities. *Review of Managerial Science, 11*(3), 543–569.

Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. *Cambridge Journal of Economics, 23*(2), 167–185.

McCann, P. (2007). Sketching out a model of innovation, face-to-face interaction and economic geography. *Spatial Economic Analysis, 2*(2), 117–134.

Moilanen, M., Østbye, S., & Woll, K. (2014). Non-R&D SMEs: External knowledge, absorptive capacity and product innovation. *Small Business Economics, 43*(2), 447–462.

Muneepeerakul, R., Lobo, J., Shutters, S. T., Goméz-Liévano, A., & Qubbaj, M. R. (2013). Urban economies and occupation space: Can they get “there” from “here”? *PLOS ONE, 8*(9), e73676.

Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal, 34*(3), 297–316.

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. The Belknap Press of Harvard University Press.

Nooteboom, B. (1994). Innovation and diffusion in small firms: Theory and evidence. *Small Business Economics, 6*(5), 327–347.

Nooteboom, B. (2000). Learning by interaction: Absorptive capacity, cognitive distance and governance. *Journal of Management and Governance, 4*(1), 69–92.

Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., & Van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy, 36*(7), 1016–1034.

OECD/Eurostat. (2018). Oslo Manual 2018: Guidelines for Collecting, Reporting, and Using Data on Innovation, 4th edition. The Measurement of Scientific, Technological, and Innovation Activities. Paris: OECD Publishing.

Rocha, F. (2018). Procurement as innovation policy and its distinguishing effects on innovative efforts of the Brazilian oil and gas suppliers. *Economics of Innovation and New Technology, 27*(8), 750–769.

Rogers, M. (2004). Networks, firm size and innovation. *Small Business Economics, 22*(2), 141–153.

Shutters, S. T., Muneepeerakul, R., & Lobo, J. (2016). Constrained pathways to a creative urban economy. *Urban Studies, 53*(16), 3439–3454.

Storper, M., & Venables, A. J. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography, 4*(4), 351–370.

Stuck, J., Broekel, T., & Revilla Diez, J. (2016). Network structures in regional innovation systems. *European Planning Studies, 24*(3), 423–442.

Torre, A., & Rallet, A. (2005). Proximity and localization. *Regional Studies, 39*(1), 47–59.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics, 113*(2), 331–360.

Whittle, A., & Kogler, D. F. (2020). Related to what? Reviewing the literature on technological relatedness: Where we are now and where can we go? *Papers in Regional Science, 99*(1), 97–113.

Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review, 93*(2), 133–138.

Zaheer, A., & Bell, G. G. (2005). Benefiting from network position: Firm capabilities, structural holes, and performance. *Strategic Management Journal, 26*(9), 809–825.

Appendix 1: Table A.1 – Variables

Variables	Type	Definition	Survey questions		
Dependent variables					
Innovation	Binary	Takes value equal to 1 when an MSE introduces an entirely or significantly improved innovation, regardless of the type, and zero otherwise.	Q29. During the last financial year (FY2019), has your establishment done any of the following types of innovation? Introduced i) entirely new products; ii) significantly improved products; iii) entirely new services; iv) significantly improved services; v) entirely new processes; vi) significantly improved processes; vii) none.		
Product innovation	Binary	Takes value equal to 1 when an MSE introduces an entirely or significantly improved product innovation, and zero otherwise.	Q29. During the last financial year (FY2019), has your establishment done any of the following types of innovation? Introduced i) entirely new products; ii) significantly improved products; iii) none.		
Process and service innovation	Binary	Takes value equal to 1 when an MSE introduces an entirely or significantly improved process or service innovation, and zero otherwise.	Q29. During the last financial year (FY2019), has your establishment done any of the following types of innovation? i) Entirely new services; ii) Significantly improved services; iii) Entirely new processes; iv) Significantly improved processes; v) None.		
Innovation novelty	Binary	Takes value equal to 1 when an MSE introduces an innovation that has the indicated degree of novelty, and zero otherwise.	Q38. In relation to the main innovation described above (Q33), was your main innovation i) only new to your business unit? ii) new to your community? iii) new to your market? iv) new to the world? v) None.		
Independent variables					
Cognitive proximity	Continuous	Indicates the levels of similar knowledge and expertise that an MSE shares with all others. It ranges from 1 to 100.	Q18. What is the main manufacturing activities you conduct in this establishment? Food products; Beverages; Tobacco products; Textiles; Wearing apparel; Leather and related products; Wood and products of wood and cork, except furniture;	Q52. What three types of skills would you say are most needed in your business? From this list, please select the three most needed skills in your business. Communication skills (e.g. internal, clients, suppliers, etc.); Marketing skills; Negotiation skills; Financial	Q56. What kind of ideas do you find most useful to exchange with other firms in your industry/trade? From this list, please select the most useful things to exchange with other firms in your industry/trade. Training; Ideas for new products or
Degree centrality	Continuous	Indicates the number of firms that an MSE is directly connected with in the knowledge network. Ranges from 1 to 100.			

Betweenness centrality	Continuou	Indicates the extent to which an MSE is located between other MSEs in the knowledge network. Ranges from 1 to 100.	Articles of straw and plaiting materials; Paper and paper products; Printing and reproduction of recorded media; Coke and refined petroleum products; Chemicals and chemical products; Pharmaceuticals, medicinal chemical and botanical products; Rubber and plastics products; Other non-metallic mineral products; Basic metals; Fabricated metal products, except machinery and equipment; Computer, electronic and optical products; Electrical equipment; Machinery and equipment; Motor vehicles, trailers and semi-trailers; Other transport equipment); Furniture; Other manufacturing; Repair and installation of machinery and equipment – except motor vehicles.	knowledge & budgetary skills; Business planning skills; Computer use; Coding and programming; Technical skills; Complex problem-solving skills; "Community skills" (knowing of the community, personal networks, etc.); Creative skills (i.e. come up with creative ideas and solutions).	services; Information about business management; Information about government support; Collective negotiations; Protecting your new ideas; None/Don't know.
Control variables					
Education	Continuous	Index indicating the educational level of the manager or owner of the MSE. Ranges from level 1 to 10.	Q4. Which of the following best describes the highest level of education completed by the owner/ manager of the business? 1) No formal school; 2) Informal schooling only; 3) Some primary school; 4) Primary school completed; 5) Some secondary school/high school; 6) Secondary/high school completed; 7) Post-secondary qualifications, other than a university (e.g. diploma or degree from a university of technology or college); 8) Some university; 9) University completed; 10) Postgraduate degree.		
R&D	Binary	Takes value equal to 1 when an MSE engages in R&D activities for innovation, and zero otherwise.	Q32 During the last financial year (FY2019), has your establishment engaged in R&D activities for innovation? (R&D refers to creative work undertaken on a systematic basis to increase the stock of knowledge and using it for innovation (including software) development).		
Skills – availability	Continuous	Indicates the extent of the skills constraint facing the MSE. Ranges from level 1 to 4.	Q46. How difficult is it for this establishment to find employees with the required skills? 1) Very difficult; 2) Fairly difficult; 3) Not very difficult; 4) Not at all difficult.		

Skills – retaining	Continuous	Indicates the capacity of an MSE to retain workers. Ranges from level 1 to 4.	Q47. How difficult is it for this establishment to retain employees? 1) Very difficult; 2) Fairly difficult; 3) Not very difficult; 4) Not at all difficult.
Skills – training	Binary	Takes value equal to 1 if employees participate in training activities (Group 1), and zero otherwise (Group 2).	Q48. What are the most important ways through which employees in this establishment become more skilled at their jobs? Please tick the most important. Group 1: Participating in training; Learning from more experienced colleagues and supervisors, for example by asking for help, through conversations, or by observing; Structured apprenticeships. Group 2: Learning by doing (trying out different ways to carry out their tasks; trial and error); Self-driven online training/Internet/YouTube.
Skills – stability	Continuous	Indicates the degree to which skills needed from workers in an MSE change over time. Ranges from level 1 to 4.	Q51. How quickly do the skills needed from the employees in this establishment change? 1) Very quickly; 2) Fairly quickly; 3) Not very quickly; 4) No change at all.
Experience	Continuous	Manager/owners' years of experience in the sector.	Q12. How many years of experience does the manager/owner have working in this field or sector?
Age of business	Continuous	Age of the business in years.	Q15. In what year did the business in this establishment begin operations?
Customer – type	Categorical	Indicates the type of customers from which an MSE gets most of its sales: individuals (baseline), business or government.	Q20. From which type of customers do you get most of your sales? Tick the most frequent type of customer for your products. Group 1: Individuals. Group 2: Small businesses; Medium businesses; Large businesses. Group 3: Government or public institutions.
Customer – local	Binary	Takes value equal to 1 if an MSE has most of its customers located in South Africa, and zero otherwise.	Q21. Where are most of your customers located?
# Suppliers	Continuous	Number of suppliers an MSE has.	Q24. How many suppliers does this establishment have? 1) None/ Don't know; 2) 1-10; 3) 11-20; 4) 21-50; 5) 51-100; 6) More than 100.
Member – business association	Binary	Takes a value equal to 1 if a MSE belongs to a business association, and zero otherwise.	Q54. Do you belong to any business association?
Co-operation	Binary	Takes a value equal to 1 if a MSE co-operates with other firms in the same industry/ sector, and zero otherwise.	Q55. Do you co-operate with other firms in the same industry/trade as you?

DSI/NRF South African Research Chair in
Industrial Development (SARChI-ID)
31 Henley Road, Auckland Park,
Johannesburg, South Africa

General enquiries:
Koketso Manyane-Dlangamandla
Email: koketsom@uj.ac.za
Tel: +27 011 559 7454