Firm performance, skill recombination and regional culture

Gianelle, Carlo

11 January 2023

Firm performance, skill recombination and regional culture

Carlo Gianelle[*]

This draft: 11 Jan 2023

Abstract

This study connects worker mobility and industry-specific skills, regional cultural characteristics and economic performance in a novel empirical way. Building on the notions of technological relatedness and unrelatedness of worker experience, we devote special attention to: (i) causal identification, and (ii) the linguistic and cultural homogeneity of the community of firms and workers under scrutiny. Contrary to several previous studies, we find that in a culturally homogeneous context, and controlling for firm heterogeneity, the unrelated combinations of skills constitute a relevant source of productivity gains. We discuss our approach in light of previous studies and derive implications for place-based policy.

Keywords: labor mobility, skill diversity, technological relatedness, firm performance

JEL Classification: D22, J62, L25, R11

[*] Senior researcher, European Commission, Directorate-General for Regional and Urban Policy, Bruxelles, Belgium. Email: carlo.gianelle@ec.europa.eu
ORCID: 0000-0003-0725-0347

Disclaimer: The views expressed here are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission.

1. Introduction

Understanding how organizations and economic systems can benefit from recombining existing knowledge to obtain new products and processes is important for growth and development. In this study, we present a novel approach to evaluating the effects of knowledge diversity – i.e. the range of different knowledge inputs potentially available for recombination – on firm performance. We analyze the mobility of skilled workers between firms within a dense, geographically bounded socio-economic community in the northeast of Italy characterized by the historical presence and current pervasive use of an unofficial language, the Venetian dialect. Such a shared communication code denotes an area of strong cultural proximity within a broader, more diversified national community. We take advantage of these unique characteristics to identify the impact of different combinations of industry-specific skills.

The spatial dimension of knowledge diffusion and interaction has proven a continuous source of inspiration for the work of and debate among economists and geographers since Alfred Marshall's (1890) seminal work, Jane Jacobs' (1969) contribution, and more recent evolutionary perspectives building upon the work of Joseph Schumpeter (1934). This was indeed the point of departure for a prolific research agenda in economic geography, focused on the measurement of diversity of knowledge inputs and its effect on innovation and growth.

The conceptual distinction introduced by Frenken et al. (2007) between technologically *related* and *unrelated variety* of economic activities and the underlying knowledge bases represents a turning point in the operationalization of a central problem in the theory of knowledge recombination, i.e. the existence of a non-linear, hump-shaped relationship between input diversity and outcomes.[1] A wealth of studies from various traditions demonstrates that diversity is a fruitful source of performance-enhancing opportunities but only up to a certain point, since too much of it can be harmful to communication and the effective integration of knowledge (Noteboom et al., 2007; Ashraf and Galor, 2013; Uzzi et al., 2013; Lee et al., 2015). The analytical solution proposed by Frenken et al. (2007) allows to capture such opposite tendencies by typifying diversity into two distinct components whose effects can be assessed separately.

An empirical 'related variety literature' emerged rapidly (Content and Frenken, 2016), focused on the effects of recombination potential on regional development (through spatially mediated externalities), and on firm performance (through direct knowledge transfer). A crucial and widely accepted result in this literature is that technologically related forms of diversity or variety of resources available as input to recombinant processes are generally good for economic performance, while technologically unrelated forms of diversity if they are not detrimental, are not conducive to growth. This proposition has been so influential that the European Union included technological relatedness as a guiding principle in the pursuit of opportunities for growth in its place-based regional innovation policy which is at the heart of the current Cohesion policy (European Commission, 2012). This position is supported also by the Organisation for Economic Co-operation and Development in its policy advice (OECD, 2013).

Especially relevant are micro-level studies on the effects of knowledge diversity embedded in labor mobility, both in light of the related variety literature (Boschma et al., 2009, 2014; Eriksson, 2011; Timmermans and Boschma, 2014) and beyond (Parrotta et al., 2014; Marino

[1] Activities are said to be technologically related or unrelated if they are respectively close or distant in a taxonomic sense when evaluated according to a given classification, e.g. the standard statistical classification of economic activities, or a metric that at least partially reflects the technological characteristics of those activities.

et al., 2016). The premise is that mobility of workers from firm to firm constitutes a major knowledge transfer channel, and matching firm and worker information allows measuring knowledge transfer directly as opposed to inferring it based, for instance, on the observation of some degree of spatiotemporal co-occurrence of events.

The evolutionary economic geography tradition to which the related variety literature belongs recognizes the influence of different forms of actor proximity on the level of coordination and the organizational pattern of economic activities. Notably, proximity can be characterized not only in purely geographical terms but also, and sometimes more importantly according to non-geographical dimensions such as technological, cultural, and institutional proximity (Boschma, 2005). This applies particularly when considering knowledge transmission through labor mobility, where linguistic and ethnic backgrounds as well as inherited systems of social norms are likely to co-determine the extent to which information and professional experience can be effectively communicated and used by others, whether individuals or organizations.

To the best of our knowledge, existing work that uses the notion of technological relatedness to assess the recombinant effect of skills on firm performance does not provide a strategy to disentangle cognitive proximity (e.g. skill diversity measured based on professional or educational experience) from other forms of proximity or heterogeneity of individual backgrounds, especially cultural backgrounds, which may influence the scope of knowledge transfer and recombination. Cultural proximity generally is not captured by geographical co-location alone, and failing to account explicitly for this type of proximity could hamper identification of the effect of cognitive skills.

To fill this gap, we propose a number of methodological innovations and test rigorously for the differential impact of technologically related and unrelated forms of diversity of worker skills. The recombinant input we consider is industry-specific knowledge accumulated on the job within a defined three-digit industry. We then evaluate the extent to which firms obtain a productivity advantage from hiring a more diverse workforce, i.e. people from a comparatively wider range of industries, distinguishing between skills acquired within the same two-digit industry of the hiring firm (technologically related) and those acquired outside (technologically unrelated). We focus on people hired to fill high-skill occupations, irrespective of their level of formal education.

The research design is based on the estimation of a firm-level production function in capital and labor, augmented by a measure of diversity of the skill profiles of recently hired workers. We apply this model to longitudinal linked employer-employee data. Our contribution to the literature is threefold: (i) isolating the effect of skill diversity from cultural dimensions affecting the transmission and recombination of worker knowledge, (iii) reducing the omitted variable bias in previous studies by properly considering the contribution of the standard production factors of capital and labor, and (iii) making explicit the identification problem involved in studying worker hiring patterns in a competitive labor market characterized by strategic matching and worker sorting.

Instrumental to our strategy is the use of the linguistic attributes of local communities as an indication of the existence of a shared communication code and broader cultural and social ties among members (Licht et al., 2007; Tabellini, 2008). More specifically, we identify in the Venetian dialect spoken in Northeastern Italy a unique case of an unofficial language that is widespread and used extensively in workplace relationships within a geographically circumscribed social group whose boundaries coincide approximately with an administrative region. Building on Falck et al. (2012), we expect that this shared communication code which is superimposed on and frequently substitutes for the official language, together with the underlying common cultural identity it reveals, will support knowledge and economic exchange among group members.

The literature studying worker skills as a recombinant input at firm level tends to ignore the identification problems stemming from the fact that worker-firm matching typically is a non-random process in competitive markets. Conceptually, we trace the possible sources of endogeneity bias to variation in firm quality which is unobserved in the data but is known at least partially to workers moving within a dense, culturally homogeneous labor market. We explicitly introduce a strategy to control for this bias by exploiting the longitudinal dimension of the data. In addition, we reduce worker self-selection into firms due to predictions of productivity shocks or the anticipation of a time-varying component of firm quality, by means of loosening the simultaneity between firm performance measures and the potentially endogenous skill diversity measures.

Our main finding is that technologically unrelated diversity of skills has a persistent positive effect on firm performance across a variety of empirical specifications. Increasing by just one category the pool of incoming skills raises total factor productivity by 0.9%, provided that those skills were acquired in sectors technologically unrelated to the industry to which the hiring firm belongs. In contrast to the standard results in the related variety literature, we find no role for diversity or magnitude of skill flows that are technologically related to the hiring firm's knowledge base.

Our analysis reveals how the effect of the recombination of individual knowledge on economic performance must be considered in close relation to the cultural and historical dimensions that influence behavior and communication among people. Using a dialect as an indication of a common cultural identity capable of supporting effective transfer and recombination of different cognitive knowledge bases is one example of how to approach this general problem. The implications of this study are particularly relevant for place-based policy. Strategies and measures aimed at fostering knowledge recombination processes involving human interaction cannot follow a generalized pattern but need to be crafted according to the specificities of places and communities.

The paper is organized as follows. Section 2 discusses the relevant contributions in different literature strands; section 3 presents the experimental context and the data; section 4 introduces the empirical strategy; section 5 reports the results; and section 6 concludes the paper.

2. Literature review

The relationship between diversity and economic performance is a key issue in the study of growth dynamics. In a seminal article, Weitzman (1998) coined the expression 'recombinant growth' to describe a Schumpeterian model where new ideas are generated through the recombination of pre-existing ones. The term diversity is used to indicate the range of ideas or knowledge inputs available for recombination. The literature on the effect of diversity on economic growth is remarkably wide and varied, and includes some studies that explore the implications for human development and the organization of society.

Typically, recombinant mechanisms are subject to a trade-off between the productivity or creativity enhancing effect of combining different information and knowledge sources, and the negative effect of integrating a variety of approaches, *modi operandi*, and languages which can reduce coordination among individuals and their joint output. This trade-off results in a hump-shaped relationship between diversity and performance which poses challenges to empirical identification and inference.

Weitzman (1998, 359) notes that "the ultimate limits to growth may lie not so much in our abilities to generate new ideas, as in our abilities to process to fruition an ever-increasing abundance of potentially fruitful ideas". Technological progress continuously improves access to information, and expands our information processing ability while reducing its cost. However, in the case of direct human interaction, two specific factors constitute additional challenges to our capacity for accessing information from people: communication failure and uncooperative or opportunistic behaviors.

To obtain relevant information from someone and to understand its value and potential use, i.e. to learn from the others, requires a shared language or communication code. Also, when information transmission has a strategic value, interpersonal trust is important for voluntary disclosure, truth telling, sustained interaction and the emergence of cooperation. Linguistic proximity enhances economic exchange as demonstrated in the literature on international trade (Melitz and Toubal, 2014). Cultural proximity, by enforcing normative social behaviors, favors mutual trust and sustains reputation mechanisms which prevent opportunism (Gellner, 2000; Tabellini, 2008), in turn supporting economic exchange (Guiso et al., 2009; Felbermeyr and Toubal, 2010).

Language is a distinctive cultural trait because its use by the speaker has the power immediately to generate in other people perception of a given cultural proximity or distance. Language has been used frequently to account for cultural diversity in economic models following Lazear's (1999a) influential contribution which points out how culture and language constitute general kinds of focal points acting as coordination devices for economic agents.[2] Accordingly, information is expected to be more effectively transferred between individuals belonging to linguistically, socially and culturally more homogeneous groups.[3] Falck et al. (2012) used historical dialects to measure cultural identity, and demonstrated that cultural similarity signaled by language positively affects economic exchange.

Furthermore, the moderating effect of culture and language on human interaction will be stronger if the relationship involves the transfer of knowledge between more diverse scientific domains, disciplines or industries. Lazear (1999b) conceptualizes the interplay of those factors in the context of multi-cultural work teams and predicts that the optimal team

---

[2] E.g. Ottaviano and Peri (2006) measured cultural diversity across U.S. cities in terms of the variety of native languages spoken by city residents. Licht et al. (2007) and Tabellini (2008) used language as an instrument for cultural variables based on the evidence that language acts as a stable constraint on cultural change; on this, see also Davis and Abdurazokzoda (2016).

[3] Lester and Piore (2004) associated the capacity of work teams to solve problems and produce innovative solutions to mutual understanding in a way that is analogous to the emergence of a community of language.

will be formed of individuals who share communication codes, are culturally homogeneous, and exhibit different and complementary skills. His predictions were confirmed empirically by Hamilton et al. (2003, 2012) based on detailed worker-level data from a single garment plant in California, and by Parrotta et al. (2014) who studied the workforce composition of Danish firms.

Research on the spatial dimension of knowledge interaction processes has looked at the link between diversity and performance mostly by applying the fundamental notions of proximity and localized knowledge spillovers.[4] Several studies provide robust evidence to support the positive role of diversity on innovation and growth in cities based on between-industry knowledge spillovers (Jacobian externalities) rather than within-industry spillovers (Marshall-Arrow-Romer externalities), see e.g. Glaeser et al. (1992) and Feldman and Audretsch (1999).

Ottaviano and Peri (2005) used workers' mother tongue to proxy for cultural identity and found that both the average wage and the employment level of U.S.-born workers are systematically and significantly higher in cities with greater cultural diversity, supporting the idea that different cultures can provide 'recombinant' inputs to production. Importantly, these effects are stronger after controlling for the degree of fluency in a common vehicular language (English) and the share of more recent and less assimilated immigrants. This shows that "the benefits of diversity are stronger when barriers between groups are lower" (Ottaviano and Peri, 2005, 327), i.e. when people share a common communication code and adhere to a common set of social norms as in Lazear (1999b). The same authors obtained similar results in a related study (Ottaviano and Peri, 2006).

[4] For a review of the concept of proximity, see Boschma (2005) and Torre and Rallet (2005); a useful review of the literature on localized knowledge spillovers can be found in Feldman (1999).

Breschi and Lissoni (2001) criticized the widespread concept of localized knowledge spillovers, and called for more precise identification of knowledge transmission mechanisms. Since then, research has focused on specific interaction processes between individuals and organizations co-located in the same area. Notably, Breschi and Lissoni (2009) found that the mobility of inventors between firms accounts for most of the co-localization of citing and cited patents' applicants, supporting the idea that knowledge is mostly transferred through labor mobility, and the observed geographical concentration of knowledge is due largely to the spatially-bounded nature of labor mobility, in line with earlier findings in Almeida and Kogut (1999).

A key innovation in studies of the impact of knowledge recombination on economic performance is the conceptual distinction between technologically related and unrelated forms of diversity introduced by Frenken et al. (2007). Thanks to this analytical device, the question of how much input diversity is optimal can be reformulated more practically in terms of which type of input diversity – related or unrelated – should be maximized.[5] This approach is at the center of a recent prolific research agenda in evolutionary economic geography.

Related variety has often been observed in association to incremental forms of innovation, while unrelated variety has been associated to breakthroughs and radical innovation (Castaldi et al., 2015). Unrelated patterns of diversification have been found to play a relatively more prominent role in more developed economies with complex product spaces (Pinheiro et al., 2022). The related variety literature builds explicitly on the notion that not all diversities are created equal; however, in contrast to Lazear's conceptualization, it has yet to consider how cultural diversity might influence communication and the recombination of cognitive skills.

[5] This problem was formulated previously in Nooteboom (1992, 1999) in terms of optimal cognitive distance.

Boschma et al. (2009) were the first to study the effects on firm productivity of different head-count measures of skill diversity, calculated based on the industry-specific experience of newly hired workers, distinguishing between the skills acquired in technologically related and unrelated industries. Using matched employer-employee data from Sweden, the authors found that hiring workers with experience in diverse industries that are technologically related to the hiring firm appears to have a positive effect on firm performance, while hiring workers with diverse skills that are technologically unrelated to the hiring firm has a negative effect. This general pattern is confirmed when focusing exclusively on inter-regional mobility, while both related and unrelated diversity measures exhibit positive effects when considering only intra-regional mobility.

Using the same data as in Boschma et al. (2009), Eriksson (2011) found very similar evidence but reported also that the negative effect of inflows of workers with technologically unrelated skills is amplified by geographical distance. Timmermans and Boschma (2014), using Danish microdata, showed that hiring workers with diverse but technologically related skills has a positive impact on firm performance countrywide, while hiring workers with diverse but technologically unrelated skills has no apparent effect, independent of geographical proximity.

None of the previously mentioned contributions investigates the interplay between individual socio-cultural characteristics and cognitive skills beyond what can be controlled for by the co-localization of firms sending and receiving worker flows. Boschma et al. (2009) assign to geographical proximity the role of controlling for ‘place-based attitudes’ but do not consider other sources of firm and worker heterogeneity. Furthermore, those studies do not account for the effect of standard production factors, capital and labor, and hence are prone to omitted variable bias. They also do not discuss problems of causal inference in the context of strategic job matching between workers and firms.

Östbring et al. (2018) provided an explicit account of the multidimensional nature of cognitive skills by distinguishing between diversity in worker educational profiles and industry experience. Again, using Swedish microdata, and looking at firms' employment stocks rather than hiring flows, they predict and seek to document the existence of trade-offs between the different skill dimensions examined. However, the evidence is inconclusive, likely due to the absence of an empirical identification strategy and bias due to omitted production factors. Östbring et al. (2017) focus on the effect of worker diversity with respect to occupations in four multi-plant, multinational Swedish firms. The results are mixed and inconclusive. The paper acknowledges the importance of social, ethnic, and linguistic diversity of workers but does not analyze these dimensions.

We extend this literature in several ways. We complement the distinction between technologically related and unrelated forms of skill diversity with a strategy to control for the cultural diversity of workers. First, we exploit the existence of a common linguistic area in the northeast of Italy coinciding with the administrative region of Veneto and the historical sovereign state of the Republic of Venice, in order to identify a culturally homogeneous population characterized by a distinctive communication code, the Venetian dialect. Second, we look at the contribution of skill diversity to total factor productivity by estimating a production function in capital and labor. Third, we explicitly tackle potential endogeneity in the estimation of worker-firm matching in order to be able to draw policy actionable conclusions.

3. Context and data

3.1 Language and identity

We focus on the region of Veneto in the northeast of Italy which in 2006, at the beginning of our period of observation, had a population of 4.702 million. The capital and largest city is Venice. The region is highly industrialized; in 2006, it was ranked 53rd among European regions for per capita gross domestic product. It has a large and differentiated “manucentric” economy characterized by mostly small and medium-sized firms very often organized in districts. The district specializations include apparel, textiles, leather and shoes, goldsmithing, machine tools, furniture, and plastics. The regional economy is export-oriented and exhibits a high level of employment compared to the Italian average (Table 1).

Veneto is notable in being delimited almost entirely by natural boundaries: the Adriatic Sea to the east, the perimeter of the drainage basin of the Piave river and the Livenza and Tagliamento rivers to the northeast, the Po river to the south and the hydrographic system of the Garda lake and Mincio river to the west. Historically, these characteristics influenced migration and contributed to maintain Veneto’s population relatively isolated, favoring the development of distinctive traits and traditions. Most relevant for the present study, Veneto’s boundaries have unique linguistic and cultural values.

The region is distinguished by the historical presence and current pervasive use of a common family of dialects, referred to collectively as the Venetian dialect. According to the classification proposed in Pellegrini (1975, 1977), Venetian is an Italo-Romance language derived directly from Vulgar Latin and belonging to the group of northern Italian dialects. The origins of the Venetian dialect can be traced back to the encounter between the ancient Venetic language – documented from the VI century BC – and Latin, before the area was integrated with the Roman Republic between the III and II centuries BC. Previous to the

Roman annexation, the population in Veneto area had established an original culture. Under the first Roman Imperator, Augustus, the Italian peninsula was organized in *regiones* defined based on ethnic, linguistic and geographical criteria, and most of actual Veneto plus the northern Adriatic coastal area, belonged to the *X Regio Venetia et Histria*, which attests to the specificity of this area.

After the fall of the Western Roman Empire in 476 AD, and the succession of barbarian rulers interrupted by a short Byzantine domination, a second fundamental phase of the socio-political history of the region began with the foundation in the year 697 AD of the *Serenissima* Republic of Venice. This lasted for more than a thousand years until its dissolution in 1797 by Napoleon Bonaparte. By the beginning of the XV century, the mainland domains of the Republic (*Stato da tera*) had extended to the whole of the Veneto region; importantly, the language commonly spoken in the capital city and its neighbouring mainland achieved the status of official language of the State. During the Republic, the Venetian language was recognized in the Mediterranean area and in continental Europe, and had a permanent influence on the vocabulary of other languages.[6]

Nowadays, Venetian is spoken across the Veneto region where it coexists with Ladino dialect in the northernmost part of the Belluno province. Its influence extends also to the Gallo-Italic dialects along the southern and western borders. According to survey data (Istat, 2007; reported in Table 1), in 2006, the Venetian dialect was spoken in family contexts by 69.9% of Veneto residents, more than 21 percentage points higher than the Italian average, with 38.9% of people using only or predominantly dialect, the highest share among Italian regions. The

[6] For instance, consider the widespread use of the informal salutation *ciao*, originally from the Venetian *s-ciào*, meaning "(I am your) slave"; also the Greek word *μαραγκός* (maragkós) and the Turkish *marangoz*, meaning carpenter, derived from the Venetian *marangone*, originally denoting a specialized ship's carpenter in the Venetian arsenal.

use of dialect with friends and acquaintances was similarly widespread, totalling 70.6% of the population, nearly 25 percentage points above the national average and first among Italian regions. Even more striking, in 2006, 15.7% of Veneto's population declared using only or predominantly dialect in exchanges with strangers compared to the national average of 5.4% and the highest percentage among Italian regions; an additional 28.7% of the population declared using both Italian and dialect with strangers.[7]

[Table 1 around here]

The high percentage of people reporting the use of both dialect and Italian in different relational contexts reflects the possibly even more widespread phenomenon of diglossia, or in the case of Veneto *dilalia* (Berruto, 1995), i.e. the ability to switch from one language to another with extreme fluidity within the same verbal relationship and even within the same sentence. According to some authors, the coexistence of Italian and Venetian, the structural similarity between these two languages, and the capacity of the speaker to use both, has contributed to the continued high level of use and comprehension of dialect in Veneto in recent times (Marcato, 2005).[8]

---

[7] In a personal communication to the authors, Alberto L., a Veneto native speaker, received at the time this article was being written, had recently been employed by an Israeli multinational company as a sales representative for Veneto, and had been told that the company was looking explicitly for individuals who could speak the Venetian dialect since this was expected to increase sales in the region.

[8] In 2007, the Veneto Regional Council officially recognized the Venetian language (Regional Law no. 8 of April 13th, 2007).

3.2 Data and measurement

Our empirical analysis is based on a unique, longitudinal, linked employer-employee dataset covering most of the private stock companies and limited companies in Veneto and their respective worker flows during the period 2006-2009.

Labor flow data come from Planet 2.1, an administrative database which contains information on worker-firm employment relationships based on data collected by territorial employment centers through firms' compulsory communications and workers' declarations (Veneto Lavoro, 2011). The database covers the universe of employment relationship in Veneto's private sector, including detailed information on individual demographics. The available data allow the construction of a daily history of workers' and firms' employment and mobility events in the region.

Firm financial data are from the AIDA database developed by the Bureau Van Dijk, which contains official balance sheet records of Italian private stock and limited companies. The available information includes detailed firm demographics, location, value added, employment, and total assets. Unique tax identification numbers allow the matching of worker information in Planet 2.1 with AIDA firm data.

Compared to the whole set of firm-worker matches obtained by merging Planet 2.1 and AIDA, the sample used in this paper is restricted to firms with both legal and main trading addresses in Veneto, which we denote Veneto firms. We measure worker skill diversity considering only Veneto natives and individuals with previous experience of working in a Veneto firm.

We restrict the research focus to the mobility of workers hired in knowledge-intensive or skilled occupations: managers, scientists, researchers, technicians, professionals, specialized administrative workers, specialized manual workers and craftsmen.[9] This provides a bigger pool of relevant workers than it would be obtained by selecting based on formal education attainment, since higher educated workers are hired mostly to fill skilled occupation vacancies, while skilled workers may not necessarily have achieved higher education. In order for workers' labor market experience to be a relevant source of knowledge, we imposed the additional constraint that they should have a minimum of six months tenure in the same industry but not necessarily in one job.

We assume that industry-specific knowledge is not instantaneously transferred when workers move into the new organization, and that the new knowledge is gradually integrated and recombined with the knowledge base of the hiring firm. Hence, we consider the flow of workers hired by firms over the three-year time window immediately preceding measurement of production output.

We construct a panel of firm-year observations in which we include only companies with at least 10 employees in each of the years 2006-2009. Our sample consists of 8,049 Veneto firms and a total of 802,123 individual worker hires over the period 2004-2009, 30% of which (242,587) are for skilled positions. The majority of these positions (89%, or 215,995 hires) are filled by Veneto natives or individuals with previous work experience in a Veneto firm.

[9] The pool of workers selected for this study includes specialized manual workers who often provide key practice-based knowledge to the firm, irrespective of their formal education attainment.

4. Empirical strategy

4.1 Skill diversity variables

The diversity of a population of individuals or set of items can be decomposed into two basic dimensions: number of different types or categories represented in the population or set, or *richness*, and relative distribution (abundance or scarcity) of individuals or items belonging to such types or categories, or *evenness* (Magurran, 1988, 2004; Maignan et al., 2003; Epping and van den Bergh, 2007; van den Heuvel and van den Bergh, 2009).

Both evenness and richness contribute positively to overall diversity. An increase in the variety of types represented in a population enriches diversity, ceteris paribus. The contribution of evenness is less intuitive. The underlying idea is that maximum diversity is obtained when no one type dominates any other, i.e. when the population is uniformly distributed across types. The more even the type distribution, the greater the diversity, ceteris paribus. We can account jointly for richness and evenness by weighting each type by the number of its representatives in the population, thereby obtaining a weighted diversity measure of the population.

As a first measure of firm level skill diversity, we evaluate the richness of industry-specific experience of newly hired workers. We do this simply by counting the number of distinct three-digit industries which the workers are hired from and which are different from the three-digit industry of the hiring firm.[10] We call this indicator skill *diversity* (denoted *div*). The assortment of skill profiles is given by the breadth of the spectrum of distinct industry experience workers bring to the hiring firm.

[10] People from the same three-digit industry as the hiring firm have an industry-specific skill endowment equivalent to or very similar to the skill base of the hiring firm, and hence according to our approach, are not primarily relevant for knowledge recombination. These flows do not contribute to the diversity measures.

The reader will notice that our focus on the composition of industry-specific skills in the flow of new hires departs from recent trends in the related-variety literature that have proposed alternative and more sophisticated ways to measure diversity. This choice is due to several reasons. On the one side, the identification of the causal effect of worker skills on firm performance requires that the taxonomy against which we evaluate diversity of worker skills be exogenous to workers' behavior. This prevents the use of information on cross-industry labor flows as a measure of revealed diversity, as for instance in Neffke and Henning (2013). On the other side, the available data do not cover the product portfolio at the plant level and hence does not allow constructing a diversity indicator based on the co-occurrence of products that belong to different industries in the portfolios of manufacturing plants, like in Neffke et al. (2011).

Moreover, the present study does not aim to propose to use the sectoral taxonomy as the best possible measure of skill diversity. It concentrates on it rather because it is an important component of individual work histories and a determinant of individual earnings, as largely shown in the literature (see for instance Neal, 1995 and Weinberg, 2001), demonstrating its relevance as a skill measure and, in turn, the suitability of the sectoral taxonomy as a diversity measure.

The second measure of diversity we consider combines information on richness and evenness of skill profiles. We use an indicator based on the structure of the Theil index which we call *weighted diversity* of skill profiles (denoted *wdiv*) and defined formally as follows for the generic firm *i*

$$wdiv_i = \sum_{j \in N^{III}} \left[\frac{x_{ij}^{III}}{X_i} \ln\left(\frac{X_i}{x_{ij}^{III}} \right) \right] \quad \forall \ j \neq j_i \qquad (1)$$

where j_i denotes the three-digit industry of firm i; N^{III} denotes both the set of three-digit industries in the economy and its cardinality, i.e. the number of distinct three-digit industries (system richness); x_{ij}^{III} is the number of workers with previous experience in the three-digit industry $j \neq j_i$ hired by firm i; X_i is the total number of workers with previous experience in any industry $j \neq j_i$ hired by firm i; and the ratio x_{ij}^{III}/X_i is the proportion of out-of-industry hires by firm i from industry j.[11]

We define as *technologically related* hires coming from the same two-digit industry as the hiring firm but from different three-digit industries to that of the hiring firm. We define as *technologically unrelated* hires from two-digit industries different from that of the hiring firm. Combining the diversity and technological relatedness of skill profiles definitions, we are able to define two sets of variables to be used in our empirical tests. Based on the index of skill diversity, we have related diversity, *reldiv*, and unrelated diversity, *unreldiv,* computed in the same way as *div* but considering only hires from technologically related and unrelated industries respectively. Based on the index of weighted skill diversity, we have weighted related diversity, *wreldiv*, and weighted unrelated diversity, *wunreldiv*, computed in the same way as *wdiv* but considering only hires from technologically related and unrelated industries respectively.

4.2 Model and estimation strategy

[11] We could argue that *div* is less sophisticated than *wdiv*, and is less complete since it overlooks the evenness dimension of diversity. At the same time, *div* has the advantage that it is more straightforward to interpret when it comes to deriving policy implications since a unit change in the indicator corresponds to a change in the skill sets of incoming workers in one three-digit industry.

The reference model for the empirical analysis is a Cobb-Douglas production function in capital and labor, calculated at firm level, and augmented by measures of worker flow characteristics capturing the diversity and the degree of technological relatedness of the skill profiles of the individuals who join the firm's workforce. The log-linear form of the model can be written as follows

$$\ln(Y_{it}) = \beta_0 + WSD_{i(t,t-2)}\beta_{WSD} + \beta_L \ln(L_{it}) + \beta_K \ln(K_{it}) + YEAR_t + SECT_i + LLM_{it} + INT_{it} + \varepsilon_{it} \quad (2)$$

The dependent variable is the value added generated by firm *i* in a given year *t*, expressed in logarithms; *WSD* is an array of diversity measures of worker skills calculated on incoming flows of workers (comprised of different, alternative combinations of *div*, *reldiv*, *unreldiv* and their weighted counterparts); *L* is total labor (workforce), and *K* is total capital (total assets, i.e. shareholders' equity plus liabilities) employed by firm *i* in year *t*, both entering the equation in logarithmic form. *YEAR* is a vector of dummies capturing time-specific productivity shocks; *SECT* is a vector of industry dummies; *LLM* is a vector of location dummies indicating the local labor market (functional region) where firms are located; *INT* is vector of year-industry interactions; and ε_{it} is an error term.

In a purely observational setting such as the one considered in this study, a potentially significant endogeneity bias derives from the fact that matching between workers and firms is typically a non-random process characterized by worker sorting and firm's strategic hiring on the basis of information that is not observed by the researcher. In general, the tighter the labor market (low unemployment), the broader the scope for worker selective behaviors. This is the case for Veneto, where in the period under scrutiny, the unemployment rate remained at historically low values (between 3.4% and 4.7%), leaving substantial scope for worker sorting. Thus, an ordinary least square (OLS) estimation of equation (2) would generate biased coefficients of the skill diversity variables, because workers moving into new jobs self-select according to their information about firm quality that is not directly observed by the researcher, and that in turn may affect firm performance. This conjecture can be expressed formally by decomposing the error term in equation (2) as follows

$$\varepsilon_{it} = \mu_i + v_{it} \tag{3}$$

where μ_i is a firm-specific productivity component indicating firm quality, that is known to the firm and perceived by the worker but is unobserved by the researcher, and v_{it} is a random term. If μ_i is correlated with the skill diversity variables *div* and *wdiv* and their related and unrelated variants included in the vector *WSD*, and if μ_i is unaccounted for in the regression equation then the OLS coefficients of β_{WSD} will be biased.

The explanation for the main mechanism underlying this type of bias is as follows. Assume that better quality firms are more productive. Assume also that prospective workers applying for job vacancies self-select according to signals about firm quality. This sorting process then alters the composition of the pool of workers applying to a given firm and could affect actual hires. More specifically, the more the information spread about firm quality is constrained and channeled by the sectoral structure of economic activities, the less random will be the sectoral composition of the applicant pool and the degree of technological relatedness of prospective workers' skills.

A well-established empirical fact in the socio-economic literature on job search is that relevant information spreads through word of mouth, and is influenced by the shape and extension of individuals' social networks (Granovetter, 1973; Cingano and Rosolia, 2012). Thus, we expect ceteris paribus, information about firm quality to spread more easily among sectors whose workers have more frequent interaction or share some common knowledge. Therefore, it is likely that the amount and quality of information about firms is negatively correlated to the technological distance between industries.[12] Better informed workers from technologically related sectors tend to crowd in on more productive firms and ignore less productive ones.

According to a slightly less intuitive logic there may be another effect at play. Information may spread beyond the localized and segmented scope of word of mouth via intermediaries. Mass media or specialized 'infomediaries' such as employment agencies (Autor, 2009; Gianelle, 2014) can spread information on firm quality that is not otherwise accessible to workers in the labor market in economic areas not connected or directly related to the firms posting vacancies. This could increase the size and heterogeneity of the pool of workers applying to better quality firms in terms of sector composition, and could include workers from unrelated, technologically distant industries. Therefore, all workers from all locations are fully informed and will crowd in on better firms.

The above mentioned phenomena are likely to occur simultaneously making the direction of the bias in model (2) difficult to predict. To provide a causal interpretation of the coefficients β_{WSD}, we first exploit the longitudinal structure of the data in order to include the vector of firm fixed effects μ_i in the estimation equations. This allows us to control for time-invariant, unobserved heterogeneity in firm quality. Second, by computing worker flows in the three-year window immediately preceding the measurement of firm performance, we considerably relax the simultaneity of the dependent variable and the potentially endogenous independent variables. This reduces the likelihood that workers self-select due to predictions of

[12] In a country such as Italy, a sector-based pattern of information spread may be induced also by the organization of the trade unions which constitute an important social-network infrastructure for the diffusion and circulation of information among workers. The Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS) database reports a trade union density of 35.2% and a strong sector focus of bargaining for Italy in the year 2011 (Visser, 2013).

productivity shocks or the anticipation of a time-varying component of firm quality. Therefore, our identifying assumption is that conditional on the cultural and linguistic homogeneity of the worker population, and after controlling for firm unobserved heterogeneity and differentiating the temporal scale of the dependent and independent variables of interest, no other relevant factor correlates simultaneously with firm performance and the skill composition of newly hired workers.

Integrating this strategy into model (2) and stating explicitly the related and unrelated components of skill diversity of worker flows (previously expressed generically as the array *WSD*), we can write the following empirical model

$$\ln(Y_{it}) = \beta_0 + \beta_{rd} reldiv_{i(t,t-2)} + \beta_{urd} unreldiv_{i,(t,t-2)} + \beta_L \ln(L_{it}) + \beta_K \ln(K_{it}) + YEAR_t + SECT_i + LLM_{it} + INT_{it} + \eta_i + \nu_{it} \quad (4)$$

The above model, together with the variant obtained for weighted diversity skill-flows variables, constitutes the most complete formulation of our empirical strategy.

5. Results

This section presents the empirical results. According to equation (4) in section 4, the regression coefficients associated to the skill diversity variables represent semi-elasticities; this implies that coefficients can be interpreted immediately as percentage changes in the outcome variable after a unit change in the corresponding skill diversity variables. Standard errors displayed in parentheses in the regression tables are clustered at firm level.

The first four columns in Table 2 report the pooled OLS estimates. In all the model specifications, labor and capital inputs are positive and significant, with the capital coefficient being about 70% of the labor coefficient which in turn is around 60% of the summed coefficients of the two inputs. If we consider only the contribution of capital and labor, the production function exhibits almost constant returns to scale, with a 1% increase in inputs leading to a 0.98% increase in output. The adjusted R-squared is close to 0.85 in all four models, showing a rather good fit. In the related variety literature that focuses on worker skills, the R-squared statistic is usually much lower, reaching 0.64 in the reference paper by Boschma et al. (2009), and revealing the possible lack of potentially important explanatory variables.[14]

[Table 2 around here]

Model 1 in Table 2 is estimated using skill diversity *div*. The coefficient of *div* is positive (0.0108) and significant. The point estimate indicates that increasing skill diversity by one

[14] The values of the adjusted R-squared statistics for the pooled OLS estimates range from 0.26 in Östbring et al. (2016) to 0.59 in Eriksson (2011), 0.66 in Timmermans and Boschma (2014), and 0.79 in Östbring et al. (2017). None of these contributions uses a production function.

unit, or equivalently, increasing the pool of insourced skills by one three-digit category increases total factor productivity by about 1.1%. Model 2 is estimated using the weighted skill diversity variable *wdiv*. The coefficient of *wdiv* is positive (0.0665) and significant. Increasing weighted skill diversity by one unit increases productivity by 6.7%.

Model 3 is estimated using the independent variables technologically related diversity *divrel* and technologically unrelated diversity *divunrel*. Our estimates reveal a 2.3% gain in productivity per one unit increase of the related diversity variable, and a 1% productivity increase per unit increase of the unrelated diversity variable. The results of the estimation of model 4 in Table 2 show that the coefficients of the related and unrelated components of weighted skill diversity are positive, statistically significant, and very similar in magnitude, implying a slightly more than 6% increase in productivity per one unit increase in either component of skill diversity.

Columns 5 to 8 in Table 2 report the respective fixed-effect estimates of the models in columns 1 to 4. In model 5, the coefficient of the variable *div* is positive (0.0087), statistically significant, and slightly smaller in magnitude than in model 1. Remarkably, skill diversity has a positive effect on firm performance even after imposing the more restrictive conditions of the fixed-effect model. This effect is confirmed by the results of the estimation of model 6, where the coefficient of the weighted skill diversity variable *wdiv* is positive (0.0303) and statistically significant, although smaller in magnitude than in model 2.

Comparison between the pooled OLS and fixed-effect models shows dramatic changes if we consider technologically related and unrelated components of skill diversity separately. The results of the estimation of model 7 show that the inclusion of firm-level fixed effects reduces the effect of related diversity *divrel* by one order of magnitude and, more important, wipes out its statistical significance. In the same model, unrelated diversity *divunrel* exhibits a positive and highly significant coefficient, implying around a 0.9% increase in total factor productivity per one unit increase of the unrelated diversity variable, only slightly lower than found for the pooled OLS estimations. Model 8 confirms this trend; the inclusion of fixed effects rules out the effect of the related component of weighted skill diversity in favor of its unrelated counterpart.

As a form of control, we conducted a separate series of estimations including in the production function a series of additional measures of incoming worker flows. The first four columns in Table 3 report the fixed-effect estimates of the four main models after including among the regressors the total number of hires of skilled workers who are Veneto natives or have previous experience of working in a Veneto firm, and who have work experience in any other sector than the three-digit industry of the hiring firm. This allows us to control for the intensity of the inflow of knowledge potentially available for recombination. Models 3 and 4 further distinguish between the intensity of incoming worker flows from technologically related and unrelated industries.

[Table 3 around here]

The skill diversity variables are still positive and significant and exhibit the same pattern found in fixed-effect estimates in Table 2. The magnitude of the coefficients is also comparable. Moreover, the role of worker flow intensity is significant and positive (0.00028) in model 2, and in model 4 (0.0023), although the magnitude of the effects is quite small. A growing number of skilled hires coming from technologically unrelated industries seems to provide more scope for knowledge recombination which in turn contributes to increase firm performance.

Columns 5 to 8 in Table 3 report the fixed-effect estimates of the four main models estimated after including among the regressors the total number of hires of skilled workers who are Veneto natives or have experience of working in a Veneto firm, and who have previous work experience in the same three-digit industry as the hiring firm. In a recombinant growth perspective, this type of workers has little to bring to the hiring firm in terms of new skills and knowledge suitable to be used for recombination; therefore, we expect the empirical effects of the corresponding variable not to be relevant. The result of the estimations confirms our expectations. The effects of the skill diversity variables in terms of overall magnitude and statistical significance are in line with previous results. Again, unrelated skill diversity emerges as the only stable source of productivity gains.

We performed a range of robustness checks based on different sample restrictions. First, we tested the role of skill diversity on a sample limited to firms employing 20 or more in each year in the sample time span. This reduced the sample used for estimation to 4,610 firms, 57% the original size. The complete results for both pooled OLS and fixed effect estimations are reported in appendix Table A1. The sign, significance and magnitude of the coefficients of the skill diversity variables are close to those reported in Table 2. According to the fixed-effect specifications, increasing technologically unrelated diversity of skills by one industry category increases productivity by approximately 0.8%, and increasing the technologically unrelated component of weighted skill diversity by one unit increases productivity by 3%.

Next, we conducted separate estimations for manufacturing and service firms. The complete fixed-effect results are reported in appendix Table A2. In the case of manufacturing firms, which account for 55% of the original sample, the sign and significance of the coefficients of the skill diversity variables are consistent with those reported in Table 2, and of similar magnitude. In the case of service firms, the signs of the coefficients of the skill diversity variables are consistent although they are of a smaller magnitude across all models and exhibit reduced significance although above the 5% level.

In summary, the results of our empirical tests show that technologically unrelated diversity of skills is a relevant source of productivity gains. The fixed-effect model rules out the contribution of the technologically related component of skill diversity. This contrasts with some of the related variety literature whose results were obtained employing models inherently more prone to omitted variable bias. It implies that in culturally and linguistically homogeneous contexts such as the one studied in this paper, technologically heterogeneous combinations of skills can be exploited (and the process to achieve them possibly need to be specifically supported), while combinations which are technologically close are already incorporated in firms' routines.

6. Conclusions

In this paper we proposed a new approach for evaluating the economic effects of knowledge recombination and we tested it empirically using longitudinal linked employer-employee data. We focused on how different combinations of industry-specific skills affect firm productivity through labor mobility, and specifically, how the diversity of such cognitive skills contributes to firm performance in possibly different ways depending on the degree of technological relatedness of the industries in which the individuals acquired their experience.

Our conceptualization of the research problem is grounded on the consideration that there are different dimensions of individual heterogeneity whose interaction may determine the scope and effectiveness of knowledge transmission and cooperation in work teams. Namely, cultural factors influencing interpersonal communication and trust can hamper the creative and productive recombination of cognitive skills developed in different technological fields and industries.

To effectively isolate the effect of skill diversity, we studied worker mobility within a socio-economic community in the northeast of Italy, corresponding geographically to the administrative region of Veneto, and characterized by the historical presence and current pervasive use of an unofficial language, the Venetian dialect. This shared communication code is a tangible sign of vital cultural ties rooted in a centuries long history of unitary history under the Republic of Venice. Workers' bonds with the regional community and its traditions through birth or previous work experience are a means of effectively controlling for otherwise unobserved cultural heterogeneity.

Besides explicitly accounting for cultural factors, our research design contributes to the existing literature by considering a firm-level production function in capital and labor, alternatively augmented by different skill diversity measures. In this setting, we discussed the potential bias due to unobserved firm heterogeneity and its effect on worker sorting. Accordingly, we explicitly formulated an identification strategy and report an exhaustive comparison between different estimation models. As a further improvement, our empirical specifications allow for direct interpretation of the magnitude of the coefficients of the variables of interest, in turn allowing quantification of possible targets for policy action. Finally, to our knowledge this is the only paper to use labor microdata for a non-Scandinavian country.

We found that technologically unrelated combinations of skills are a relevant source of productivity gains across a variety of empirical specifications; our identification strategy clearly and systematically rules out effects due to technologically related combinations of skills. Our results differ substantially from those reported in previous works in the related variety literature, and challenge a widely accepted interpretation of technologically related diversity as a general evolutionary principle governing economic growth.[15]

Our approach and the evidence we provide support a novel set of policy implications. Namely, in contexts where people share cultural traits and communication codes, or more generally, contexts characterized by widely accepted systems of social norms whose adherence can be clearly signaled to other members of the community, training programs, education policies, and labor market policies in support of mobility between industries maximize their potential if aimed at establishing or multiplying mobility channels between

[15] In a personal communication to the authors, Mikel N.A., a scholar who did research on the Basque Country region of Spain using the conventional related variety framework lenses, told us that, based on labor mobility flows between functional regions, he had been unable to find evidence compatible with the predictions of the related variety approach. This led to the research project being abandoned. Given the cohesion of the Basque community from a cultural and linguistic point of view, the approach proposed in this paper would have produced different expectations, perhaps closer to the actual empirical evidence.

activities and firms characterized by different technological bases. Similarly, in socially cohesive contexts industrial policies, and especially those implemented at the sub-national scale, ought preferentially to support knowledge transfer according to technologically unrelated patterns across industries.

Our results have potential significance for place-based policies especially related to research and innovation.[16] Such policies should adapt their means and rationale not only to the socio-economic conditions of specific territories but also and fundamentally to the cultural and linguistic characteristics of the communities living in those territories, and to the local presence of institutions that effectively can support interpersonal communication and trust.

[16] E.g. European Union Cohesion policy supported initiatives for the funding periods 2014-2020 and 2021-2027 where the so called "smart specialization" approach provides a place-based paradigm for public intervention aimed to support research and innovation.

References

Almeida, P., & Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science* 45(7), 905-917.

Ashraf, Q., & Galor, O. (2013). The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review* 103(1), 1-46.

Autor, D.H. (2009). (ed.) *Studies of Labor Market Intermediation*. Chicago: University of Chicago Press.

Berruto, G. (1995). *Fondamenti di sociolinguistica*. Roma and Bari: Laterza.

Boschma, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies* 39(1), 61-74.

Boschma, R., Eriksson, R.H., & Lindgren, U. (2009). How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography* 9(2), 169-190.

Boschma, R., Eriksson, R.H., & Lindgren, U. (2014). Labour Market Externalities and Regional Growth in Sweden: The Importance of Labour Mobility between Skill-related Industries. *Regional Studies* 48(10), 1669-1690.

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

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography* 9(4), 439-468.

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.

Cingano, F., & Rosolia, A. (2012). People I Know: Job Search and Social Networks. *Journal of Labor Economics* 30(2), 291-332.

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

Davis, L.S., & Abdurazokzoda, F. (2016). Language, culture and institutions: Evidence from a new linguistic dataset. *Journal of Comparative Economics* 44(3), 541-56.

Epping, F.V., & van den Bergh, J.C.J.M. (2007). Ecological theories and indicators in economic models of biodiversity loss and conservation: A critical review. *Ecological Economics* 61, 284-293.

Eriksson, R.H. (2011). Localized Spillovers and Knowledge Flows: How Does Proximity Influence the Performance of Plants?. *Economic Geography* 87(2), 127-152.

European Commission (2012). *Guide to Research and Innovation Strategies for Smart Specialisation (RIS3)*. Luxembourg: Publication Office of the European Union.

Falck, O., Heblich, S., Lameli, A., & Südekum, J. (2012). Dialects, cultural identity, and economic exchange. *Journal of Urban Economics* 72, 225-239.

Felbermayr G.J., & Toubal, F. (2010). Cultural proximity and trade. *European Economic Review* 54(2), 279-293.

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

Feldman, M.P., & Audretsch, D. (1999). Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition. *European Economic Review* 43, 409-429.

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

Gellner, E. (2000). Trust, Cohesion, and the Social Order. In Gambetta, D. (ed.). *Trust: Making and Breaking Cooperative Relations*. Electronic edition, Department of Sociology, University of Oxford.

Gianelle, C. (2014). Labour market intermediaries make the world smaller. *Journal of Evolutionary Economics* 24(5), 951-981.

Glaeser, E., Kallal, H., Scheinkman, J., & Sheifler, A. (1992). Growth of Cities. *Journal of Political Economy* 100, 1126-52.

Granovetter, M.S. (1973). The Strength of Weak Ties. *American Journal of Sociology* 78, 1360-1380.

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural Biases in Economic Exchange?. *Quarterly Journal of Economics* 124(3), 1095-1131.

Hamilton, B.H., Nickerson, J.A., & Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy* 111(3), 465-497.

Hamilton, B.H., Nickerson, J.A., & Owan, H. (2012). Diversity and Productivity in Production Teams. In Bryson, A. (ed.). *Advances in the Economic Analysis of Participatory and Labor-Managed Firms (Volume 13)*. Bingley: Emerald Group Publishing Limited.

Istat (2007). La lingua italiana, i dialetti e le lingue straniere. Statistiche in breve, Istituto Nazionale di Statistica. Rome, 20 April.

Istat (2008). Struttura e dimensione delle imprese, Archivio Statistico delle Imprese Attive (ASIA). Statistiche in breve, Istituto Nazionale di Statistica. Rome, 24 July.

Jacobs, J. (1969). *The Economy of Cities*. NewYork: Random House.

Lazear, E.P. (1999a). Culture and Language. *Journal of Political Economy* 107(S6), S95-S126.

Lazear, E.P. (1999b). Globalisation and the Market for Team-Mates. *The Economic Journal* 109(454), C15-C40.

Lee, Y.-N., Walsh, J.P., & Wang, J. (2015). Creativity in scientific teams: Unpacking novelty and impact. *Research Policy* 44, 684-697.

Lester, R.K., & Piore, M.J. (2004). *Innovation: The Missing Dimension*. Cambridge: Harvard University press.

Licht, A.N., Goldschmidt, C., & Schwartz, S.H. (2007). Culture rules: The foundations of the rule of law and other norms of governance. *Journal of Comparative Economics* 35(4), 659-688.

Magurran, A.E. (1988). *Ecological diversity and its measurement*. London: Croom Helm.

Magurran, A.E. (2004). *Measuring biological diversity*. Oxford: Blackwell Publishing.

Maignan, C., Ottaviano, G.I.P., Pineli, D., & Rullani, F. (2003). Bio-ecological diversity vs. socio-economic diversity: a comparison of existing measures. FEEM Working Paper no. 13.2003. Fondazione Eni Enrico Mattei, January 2003.

Marcato, G. (ed.) (2005). *Lingue e dialetti nel Veneto. Volume 3*. Padova: Unipress.

Marino, M., Parrotta, P., & Pozzoli, D. (2016). Educational diversity and knowledge transfers via inter-firm labor mobility. *Journal of Economic Behavior & Organization* 123, 168-183.

Marshall, A. (1890). *Principles of Economics*. Reprinted: London: Palgrave Macmillan, 2013.

Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics* 93, 351-363.

Neffke, F., Henning, M., & Boschma, R. (2011). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography* 87(3), 237–265.

Neffke, F., & Henning, M. (2013). Skill Relatedness and Firm Diversification. *Strategie Management Journal* 34, 297-316.

Neal, D. (1995). Industry-Specific Human Capital: Evidence from Displaced Workers. *Journal of Labor Economics* 13(4), 653-677.

Nooteboom, B. (1992). Towards a dynamic theory of transactions. *Journal of Evolutionary Economics* 2, 281–299.

Nooteboom, B. (1999). Innovation and inter-firm linkages: new implications for policy. *Research Policy* 28, 793-805.

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

OECD (2013). *Innovation-driven Growth in Regions: The Role of Smart Specialisation*. Paris: OECD Publications.

Östbring, L., Eriksson, R.H., & Lindgren, U. (2018). Relatedness through experience: On the importance of collected worker experiences for plant performance. *Papers in Regional Science* 97(3), 501-519.

Östbring, L., Eriksson, R.H., & Lindgren, U. (2017). Labour mobility and organisational proximity: routines as supporting mechanisms for variety, skill integration and productivity. *Industry and innovation* 24(8), 775-794.

Ottaviano, G.I.P., & Peri, G. (2005). Cities and cultures. *Journal of Urban Economics* 58(2), 304-337.

Ottaviano, G.I.P., & Peri, G. (2006). The economic value of cultural diversity: evidence from US cities. *Journal of Economic Geography* 6(1), 9-44.

Parrotta, P., Pozzoli, D., & Pytlikova, M. (2014). Labor diversity and firm productivity, *European Economic Review* 66, 144-179.

Pellegrini, G.B. (1975). I cinque sistemi dell'italo-romanzo. In Pellegrini, G.B. (ed.). *Saggi di linguistica italiana. Storia, struttura, società*. Torino: Boringhieri.

Pellegrini, G.B. (1977). *Carta dei dialetti d'Italia*. Pisa: Pacini.

Pinheiro, F.L. ,Hartmann, D., Boschma, R. & Hidalgo, C.A. (2021). The time and frequency of unrelated diversification. *Research Policy* 51(8), 104323.

Schumpeter, J.A. (1934). *The Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*. Reprint: New Brunswick and London: Transaction Publishers, 2012.

Tabellini, G. (2008). Institutions and Culture. *Journal of the European Economic Association* 6(2-3), 255-294.

Timmermans, B., & Boschma, R. (2014). The effect of intra- and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows. *Journal of Economic Geography* 14(2), 289-311.

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

Uzzi, B., Mukerjee, S., Stringer, M., & Jones, B.F. (2013). Atypical Combinations and Scientific Impact. *Science* 342, 268-472.

van den Heuvel, S.T.A., & van den Bergh, J.C.J.M. (2009). Multilevel assessment of diversity, innovation and selection in the solar photovoltaic industry. *Structural Change and Economic Dynamics* 20, 50-60.

Veneto Lavoro (2011). PLANET 2.0 e successive. Note e specifiche guida. Documento Interno N. 50. Veneto Lavoro Osservatorio & Ricerca. Mestre-Venezia, January.

Visser, J. (2013). Data Base on Institutional Characteristics of Trade Unions, Wage Setting, State Intervention and Social Pacts (ICTWSS) version 4.0. Amsterdam Institute for Advanced Labour Studies (AIAS), University of Amsterdam, April 2013.

Weinberg, B.A. (2001). Long-Term Wage Fluctuations with Industry-Specific Human Capital. *Journal of Labor Economics* 19(1), 231-264.

Weitzman, M.L. (1998). Recombinant Growth. *Quarterly Journal of Economics* 113(2), 331-360.

Tables

Table 1 – Veneto characteristics, 2006

	Veneto (national share)	Italy
Population (thousands people)	4,702 (8.1%)	58.064
GDP (million EUR, current prices)	140,576 (9.4%)	1,493,031
Export (million EUR)	46,284 (13.9%)	332,013
Employment in firms (thousands people):		
- *Total*	*1,685.8 (9.8%)*	*17,116.8*
- *Industry*	*605.6 (12.8%)*	*4,730.3*
- *Share of industry employment*	*35.9%*	*27.6%*
Unemployment rate	4.1%	6.8%
Average firm size (number of workers):		
- *Total*	*4.2*	*3.9*
- *Industry*	*10.6*	*9.1*
People using only or prevalently dialect:		
- *With family members*	*38.9%*	*16.0%*
- *With friends and acquaintances*	*37.3%*	*13.2%*
- *With strangers*	*15.7%*	*5.4%*
People using both Italian and dialect:		
- *With family members*	*31.0%*	*32.5%*
- *With friends and acquaintances*	*33.3%*	*32.8%*
- *With strangers*	*28.7%*	*19.0%*

Sources: database i.Stat (available at: http://dati.istat.it/); Istat (2007, 2008).

Table 2 – Effects of skill diversity

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) FE	(6) FE	(7) FE	(8) FE
div	0.0108^{***}				0.0087^{***}			
	(0.0014)				(0.0016)			
wdiv		0.0665^{***}				0.0303^{***}		
		(0.0062)				(0.0055)		
divrel			0.0231^{***}				0.0031	
			(0.0065)				(0.0060)	
divunrel			0.0101^{***}				0.0089^{***}	
			(0.0015)				(0.0016)	
wdivrel				0.0621^{**}				0.0124
				(0.0202)				(0.0195)
wdivunrel				0.0634^{***}				0.0305^{***}
				(0.0064)				(0.0057)
log(labour)	0.5792^{***}	0.5796^{***}	0.5793^{***}	0.5794^{***}	0.3940^{***}	0.3995^{***}	0.3943^{***}	0.3999^{***}
	(0.0109)	(0.0105)	(0.0109)	(0.0106)	(0.0240)	(0.0243)	(0.0240)	(0.0242)
log(capital)	0.4008^{***}	0.3992^{***}	0.4006^{***}	0.3992^{***}	0.4243^{***}	0.4260^{***}	0.4244^{***}	0.4259^{***}
	(0.0068)	(0.0068)	(0.0068)	(0.0068)	(0.0174)	(0.0177)	(0.0174)	(0.0177)
constant	1.7119^{***}	1.7229^{***}	1.7127^{***}	1.7222^{***}	2.3517^{***}	2.3215^{***}	2.3501^{***}	2.3219^{***}
	(0.1496)	(0.1490)	(0.1495)	(0.1490)	(0.1643)	(0.1692)	(0.1643)	(0.1686)
Adj. R-sqr	0.8483	0.8487	0.8483	0.8486	0.2183	0.2172	0.2183	0.2172
N. of obs.	31898	31898	31898	31898	31898	31898	31898	31898
N. of firms	8049	8049	8049	8049	8049	8049	8049	8049

Dependent variable: log value added. Standard errors in parentheses clustered by firm. Columns 1 to 4 report OLS estimates based on alternative specifications of skill diversity variables. Columns 5 to 8 report firm-level fixed-effect estimates based on alternative specifications of skill diversity variables. All estimations include year dummies, industry dummies, LLM dummies, and year-industry interactions. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3 – Effects of skill diversity controlling for incoming worker flows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	FE	FE	FE	FE	FE	FE	FE
div	0.0075***				0.0079***			
	(0.0022)				(0.0014)			
wdiv		0.0203***				0.0289***		
		(0.0056)				(0.0055)		
divrel			-0.0042				0.0026	
			(0.0074)				(0.0060)	
divunrel			0.0097***				0.0082***	
			(0.0021)				(0.0015)	
wdivrel				-0.0041				0.0091
				(0.0211)				(0.0187)
wdivunrel				0.0216***				0.0290***
				(0.0057)				(0.0056)
total recomb hires	0.0008	0.0028***						
	(0.0011)	(0.0008)						
related recomb hires			0.0052	0.0050				
			(0.0036)	(0.0033)				
unrelated recomb hires			-0.0006	0.0023**				
			(0.0009)	(0.0008)				
same sector inflow					0.0008	0.0010	0.0008	0.0010
					(0.0006)	(0.0006)	(0.0006)	(0.0006)
log(labour)	0.3938***	0.3943***	0.3941***	0.3951***	0.3928***	0.3963***	0.3931***	0.3968***
	(0.0240)	(0.0243)	(0.0240)	(0.0242)	(0.0238)	(0.0239)	(0.0238)	(0.0238)
log(capital)	0.4242***	0.4245***	0.4240***	0.4244***	0.4231***	0.4240***	0.4232***	0.4239***
	(0.0174)	(0.0176)	(0.0174)	(0.0176)	(0.0169)	(0.0170)	(0.0169)	(0.0170)
constant	2.3531***	2.3490***	2.3539***	2.3479***	2.3654***	2.3468***	2.3639***	2.3466***
	(0.1644)	(0.1670)	(0.1642)	(0.1669)	(0.1576)	(0.1584)	(0.1577)	(0.1583)
Adj. R-sqr	0.2183	0.2181	0.2186	0.2181	0.2188	0.2181	0.2188	0.2181
N. of obs.	31898	31898	31898	31898	31898	31898	31898	31898
N. of firms	8049	8049	8049	8049	8049	8049	8049	8049

Dependent variable: log value added. Standard errors in parentheses clustered by firm. Columns 1 to 4 report firm-level fixed-effect estimates based on alternative specifications of skill diversity variables, controlling for the total number of hires from any other sector than the three-digit industry of the hiring firm. Columns 5 to 8 report firm-level fixed-effect estimates based on alternative specifications of skill diversity variables, controlling for the number of new hires from the same 3-digit industry of the hiring firm. All estimations include year dummies, industry dummies, LLM dummies, and year-industry interactions. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix

Table A1 – Effects of skill diversity, only firms with at least 20 employees

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	FE	FE	FE	FE
div	0.0099***				0.0081***			
	(0.0016)				(0.0017)			
wdiv		0.0674***				0.0309***		
		(0.0077)				(0.0068)		
divrel			0.0211**				0.0064	
			(0.0074)				(0.0073)	
divunrel			0.0092***				0.0081***	
			(0.0017)				(0.0018)	
wdivrel				0.0558*				0.0222
				(0.0226)				(0.0222)
wdivunrel				0.0650***				0.0304***
				(0.0078)				(0.0070)
log(labour)	0.5492***	0.5526***	0.5492***	0.5515***	0.4248***	0.4383***	0.4248***	0.4378***
	(0.0154)	(0.0146)	(0.0154)	(0.0147)	(0.0370)	(0.0376)	(0.0369)	(0.0374)
log(capital)	0.4276***	0.4253***	0.4274***	0.4253***	0.4432***	0.4450***	0.4433***	0.4448***
	(0.0096)	(0.0095)	(0.0096)	(0.0095)	(0.0255)	(0.0260)	(0.0255)	(0.0259)
constant	1.5810***	1.5884***	1.5816***	1.5897***	2.1640***	2.0964***	2.1635***	2.1013***
	(0.1968)	(0.1952)	(0.1966)	(0.1953)	(0.2644)	(0.2749)	(0.2645)	(0.2734)
Adj. R-sqr	0.8295	0.8301	0.8296	0.8301	0.2257	0.2243	0.2257	0.2243
N. of obs.	18264	18264	18264	18264	18264	18264	18264	18264
N. of firms	4610	4610	4610	4610	4610	4610	4610	4610

Dependent variable: log value added. Standard errors in parentheses clustered by firm. Columns 1 to 4 report OLS estimates based on alternative specifications of skill diversity variables. Columns 5 to 8 report firm-level fixed-effect estimates based on alternative specifications of skill diversity variables. All estimations include year dummies, industry dummies, LLM dummies, and year-industry interactions. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A2 – Effects of skill diversity, manufacturing and service firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	FE	FE	FE	FE	FE	FE	FE
	M	M	M	M	S	S	S	S
div	0.0092^{***}				0.0056^{***}			
	(0.0022)				(0.0016)			
wdiv		0.0306^{***}				0.0239^{*}		
		(0.0071)				(0.0100)		
divrel			-0.0040				0.0093	
			(0.0074)				(0.0102)	
divunrel			0.0100^{***}				0.0055^{**}	
			(0.0023)				(0.0017)	
wdivrel				0.0086				0.0069
				(0.0236)				(0.0277)
wdivunrel				0.0300^{***}				0.0248^{*}
				(0.0072)				(0.0102)
log(labour)	0.4577^{***}	0.4669^{***}	0.4588^{***}	0.4677^{***}	0.3314^{***}	0.3318^{***}	0.3311^{***}	0.3320^{***}
	(0.0371)	(0.0378)	(0.0370)	(0.0376)	(0.0369)	(0.0370)	(0.0369)	(0.0369)
log(capital)	0.4600^{***}	0.4627^{***}	0.4602^{***}	0.4625^{***}	0.3979^{***}	0.3981^{***}	0.3978^{***}	0.3982^{***}
	(0.0229)	(0.0239)	(0.0230)	(0.0239)	(0.0317)	(0.0316)	(0.0317)	(0.0316)
constant	1.8406^{***}	1.7904^{***}	1.8354^{***}	1.7904^{***}	2.7872^{***}	2.7845^{***}	2.7891^{***}	2.7837^{***}
	(0.2306)	(0.2447)	(0.2309)	(0.2436)	(0.2836)	(0.2824)	(0.2837)	(0.2824)
Adj. R-sqr	0.2207	0.2193	0.2209	0.2192	0.2090	0.2087	0.2090	0.2087
N. of obs.	17594	17594	17594	17594	10693	10693	10693	10693
N. of firms	4435	4435	4435	4435	2703	2703	2703	2703

Dependent variable: log value added. Standard errors in parentheses clustered by firm. Columns 1 to 4 report firm-level fixed-effect estimates based on alternative specifications of skill diversity variables, considering only manufacturing firms. Columns 5 to 8 report firm-level fixed-effect estimates based on alternative specifications of skill diversity variables, considering only service firms. All estimations include year dummies, industry dummies, LLM dummies, and year-industry interactions. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Statements and Declarations

Compliance with Ethical Standards

The manuscript is an original work and is not submitted for publication in any other journal. This research does not involve human participants or animals.

Competing Interests and Funding

The author has no financial or non-financial conflicts of interest to disclose.

Data Availability

The datasets analyzed during the current study are not publicly available as they contain proprietary information that the authors acquired through a license (AIDA database developed by the Bureau Van Dijk), and sensitive personal information about work lives the access to which was granted to the author on an individual basis for research purposes (Planet 2.1 database developed by Veneto Lavoro and available in Italian only). Information on how to obtain the datasets and reproduce the analysis is available from the corresponding author on request.