



Does language prevent policy take-up? Evidence from the Italian Start-up Act

Michele Cantarella^{a,b,*}, Nicolò Fraccaroli^{c,d}, Roberto Volpe^e

^a Scuola IMT Alti Studi Lucca, Italy

^b University of Helsinki, Finland

^c Brown University, United States of America

^d World Bank, United States of America

^e LUISS Guido Carli, Italy

ARTICLE INFO

JEL classification:

M13

Z13

D83

O38

O32

Keywords:

Ethnolinguistic asymmetries

Start-ups

Innovation policy

Network embeddedness

ABSTRACT

Does ethnolinguistic diversity prevent policy adoption? The implementation of the Italian Start-up Act of 2012 in the bilingual (German and Italian) region of Trentino-Alto Adige offers the ideal setting to investigate this question. The Act sets up a scheme of benefits which young firms can access by registering as “innovative start-ups” on a voluntary basis. We find that policy take-up has been persistently lower in areas of the region with more German speakers, as local firms with German-named administrators are less likely to register as start-ups than firms with Italian-named ones. These findings are robust to firm characteristics and regional heterogeneity and are also visible within mixed-language municipalities. Furthermore, text analysis on press sources suggests that this national policy was much more extensively covered in the Italian-language local media, while a survey of local residents indicates that German speakers have lower knowledge of national policies unless they are embedded in multilingual networks.

1. Introduction

Does language prevent policy take-up? Existing works highlighted that linguistic cleavages affect the supply of policies by lowering the provision of public goods (Beach and Jones, 2017; Alesina et al., 2003; La Porta et al., 1999), hindering redistribution (Desmet et al., 2012; Dahlberg et al., 2012) and growth (Alesina et al., 2003), or worsening government performance (La Porta et al., 1999).¹ However, evidence that linguistic differences pose specifically a barrier to uptake of policies is scarce. Does access to a policy vary across ethnolinguistic groups in the same country? Are ethnolinguistic minorities less likely to seek access to the benefits of a policy, let alone be aware of certain policies?

In this paper, we investigate these questions based on the case of the Italian *Start-up Act*, a policy introduced in 2012 with the aim of boosting investment in innovation among start-up firms. The policy, which is still in force, provides a set of economic and legal benefits

to eligible firms, which are recorded as ‘innovative start-ups’. These benefits include tax incentives, easier access to public credit guarantees, and more flexible legal provisions (fewer restrictions on contracts, protection from bankruptcy, etc.). This specific policy suits our research question particularly well because it is a widely advertised national industrial policy based on voluntary access: eligible firms that are aware of the policy must submit a request to be registered as innovative start-ups and enjoy its benefits.

We focus on the region of Trentino-Alto Adige/South Tyrol.² This Italian region located on the border with Austria provides a unique setting to study how linguistic cleavages affect policy access. A historical linguistic divide characterises the inhabitants of the region who, for the rest, are similar in terms of entrepreneurship and socio-economic opportunities: while they are all Italian citizens, some of them speak Italian, whereas a considerable proportion (around 35%)³ speaks German as their main language. Our key hypothesis is that

* Corresponding author at: Scuola IMT Alti Studi Lucca, Italy.

E-mail address: michele.cantarella@imtlucca.it (M. Cantarella).

¹ For a review, see Alesina and La Ferrara (2005).

² The official name of this region is ‘Trentino-Alto Adige/Südtirol’, acknowledging the equal footing the German language has with Italian in the autonomous province of Bolzano (Bozen in German). This province is normally called, in English language sources, ‘South Tyrol’, a translation of Südtirol. For clarity and concision only, we will use this name throughout the paper.

³ Official statistics on linguistic groups are provided for South Tyrol, while Trentino is almost unilingually Italian. The German-speaking group accounts for 69.41% of the total population of the province of Alto Adige/South Tyrol. Source: ASTAT’s population census of 2011. Available at: <https://redas.services.siaig.it/redasArticlesAttachment?attachId=562911> (last accessed on 26 January 2022).

ethnolinguistic diversity can hamper the spread of information about the policy within entrepreneurial networks.

In fact, eligible companies with German-speaking management appear to be far less likely to apply for the status of innovative start-ups than those with Italian-speaking managers who share similar characteristics. While many of these differences could be attributed to institutional, industrial and policy differences within the region, we argue that this might not always be the case. On top of these territorial factors, language group assignment might also affect embeddedness to linguistic networks, giving rise to informational asymmetries.

To test this hypothesis, we follow three distinct approaches, gaining comprehensive insights into the policy's coverage in the press, residents' awareness, and firm adoption patterns, shedding light on the impact of linguistic factors in policy dissemination and implementation.

Firstly, study the news coverage of this policy using text analysis methods within the corpus of articles from in-press and online local news outlets from 2010 to 2019. Our analysis of policy coverage from local press sources suggests that informational asymmetries do arise in the press, as German-language sources are far less likely to mention the policy when compared to the Italian-language press.

Secondly, we study residents' awareness of a bundle of five government policies, including the Start-up Act, making use of a recent survey of ~700 respondents for a total of ~3500 pooled observations. We find that individuals whose native language is German are less knowledgeable about national policies than their Italian-speaking counterparts. These results are robust to self-reported language proficiency and a variety of location factors. We find that the only factor that eliminates this knowledge gap is a sustained level of interaction with Italian speakers.

Finally, we analyze the take-up of the policy on a sample of ~20,000 firms based in the region to study the effect of ethnolinguistic assignment on the take-up of the Start-up Act. We identify a firm's membership to a language group by looking at the etymological origin of the surname of its CEO and its Board of Directors, developing a simple text-based method to classify over 55,000 board members' surnames as having Germanic (German-speaking) or Romance (Italian-speaking) roots. Hence, we exploit the etymological linguistic origins of the managers' surnames as an exogenous source of variation to predict the likelihood that a firm accesses the policy. The fact that surnames are outside the control of individuals means that the surname distribution is only geographically determined,⁴ so conditional independence is easier to obtain by controlling for space–time variation in several ways. Our findings suggest that linguistic asymmetries persist in the registration behaviour of firms regardless of regional heterogeneity or policy shocks, even within the same municipalities or industries. Moreover, once these factors are controlled for, we find no differences in performance or innovation propensity between the two linguistic groups, as we show that eligible firms with German-speaking CEOs (or boards) performed just as well as their counterparts with Italian-speaking CEOs in the years preceding the introduction of the policy.

These findings suggest that some differences in policy adoption cannot be explained by economic and policy differences between municipalities but, rather, can be attributed to ethnolinguistic assignment only. While the immediate cause might be disinterest, detachment from the national sphere, firm-level variation in informal institutions, or, as our results suggest, linguistic network embeddedness, what matters is that all these factors are ultimately determined by language assignment and coalesce into a policy take-up, and awareness, bias. Informal talks with local actors in the entrepreneurial system confirmed this view.

Our work contributes to several strands of the literature. First and foremost, we complement the broader literature of cultural economics

(e.g., Desmet and Wacziarg, 2020; Alesina and Giuliano, 2015; Guiso et al., 2015; Guiso et al., 2008), and in particular those studies that focus on the economic costs that stem from linguistic differences. While prior research showed how language diversity shapes economic behaviour (Montalvo and Reynal-Querol, 2021; Wang and Steiner, 2015; Falk et al., 2018; Galor et al., 2020; Accetturo et al., 2023, 2019) and the supply of policies and institutions (Bazzi et al., 2019; Beach and Jones, 2017; Desmet et al., 2016; Spolaore and Wacziarg, 2016; Desmet et al., 2012; Dahlberg et al., 2012; Alesina et al., 2003; Alesina and La Ferrara, 2005; La Porta et al., 1999), evidence on how linguistic cleavages affect policy access is less widespread. Some studies, such as Chen (2013) and, on Trentino-Alto Adige specifically, Sutter et al. (2018) have found that languages can affect intertemporal preferences, but there is little evidence on whether languages can actually influence entrepreneurship activity. Our research also ties with the literature on how network embeddedness (Granovetter, 1973) relates to ethnicity (Larson and Lewis, 2017) and entrepreneurship (Gallo and Plunket, 2020; Andersen, 2013; Gilsing et al., 2008; Cooke et al., 2005).

Second, our evidence contributes to the scholarship of innovation economics that studies the role of start-ups and the policies supporting them. Start-ups are considered engines of economic growth thanks to their high potential for innovation (Andrews et al., 2014; Baumol, 2004; Acs and Audretsch, 1987) and important sources of job creation (Buldyrev et al., 2020; 2007; Calvino et al., 2015; Haltiwanger et al., 2013; Criscuolo et al., 2014). As a response to the decline in firm dynamism (Decker et al., 2016) and trends in automation and offshoring (Autor, 2013), governments of advanced economies increasingly engaged in the design of policies to promote the formation and growth of young innovative firms (Schot and Steinmueller, 2018; OECD, 2020c; Breschi et al., 2018). These actions led to the proliferation of start-ups, which today account for about 20% of employment across OECD countries (OECD, 2020c) and for almost one-half of productivity growth in the United States (Klenow and Li, 2020). As small firms might be more susceptible to the costs arising from uneven information flows, which have been shown to affect innovation (Murray et al., 2016), understanding which factors, including language assignment and its mediators, contribute to policy take-up can help enhance the delivery of public policies directed towards young innovative enterprises.

The remainder of this paper is structured as follows. The next section provides a background on the Start-up Act policy framework and on the linguistic divide in the Trentino-Alto Adige region, framing it within the network embeddedness discourse. Section 3 describes the data we use and their sources. Section 4 presents the results from our text analysis exercise. Policy awareness results are shown in Section 5, while policy adoption results are offered in Section 6. Section 7 concludes. The Appendix describes our data sources and text mining methods in detail, offers additional robustness checks, offers additional qualitative interview evidence, and provides a framework for exploiting these ethnolinguistic asymmetries for policy evaluation.

2. Background and motivation

2.1. The Start-up Act

This paper studies take-up of a package of benefits created for young innovative companies ('*Start-up innovativa*' in Italian, i.e. Innovative Start-up) based in Italy, which was launched in late 2012 and has been functional since early 2013, colloquially called the 'Start-up Act'.⁵ As stated in its preamble, the goal of the Start-up Act is to

⁴ The Italian Civil Code (article 143 bis) also requires all married women to keep their maiden name unless otherwise decided (which is usually not the case for German-speaking countries), making the Italian setting well suited for this identification strategy.

⁵ The full text of the Act can be browsed on Italy's Official Gazette, at: https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2012-12-18&atto.codiceRedazionale=12A13277 (Last accessed: 12 July 2023).

promote ‘sustainable growth, technological development, innovative entrepreneurship and youth employment’, and thereby contribute to social mobility and the attraction of innovative firms and capital in the country.

Innovative start-ups are assigned several facilitations that aim to benefit all stages of start-ups’ life-cycle, from birth to maturity (MISE, 2019), including several regulatory simplifications, such as exemptions from duties and fees, and corporate governance amendments, such as a waiver from standard bankruptcy proceedings. Between 2016 and 2021, a simplified form of company incorporation was also in place. On the financing side, a public guarantee is applied automatically to all approved loan applications by innovative start-ups. Unlike most companies, the public guarantee is applied at no cost to the entrepreneurs and with no additional due diligence carried out beyond the checks already done by the bank issuing the loan. The policy also supports equity financing via tax breaks for seed- and early-stage investors.

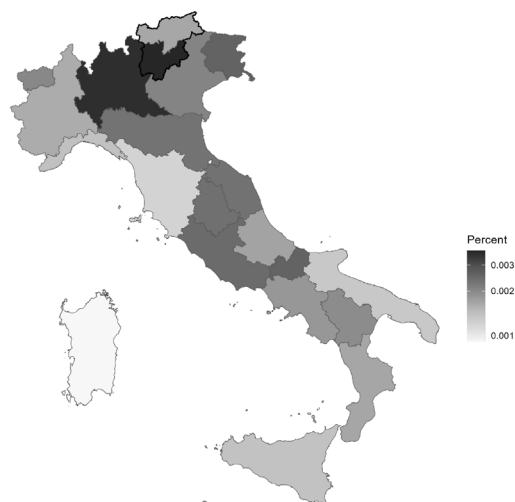
To be eligible for the policy, start-ups need to be limited liability companies, less than five years old,⁶ not be publicly listed, and report an annual turnover lower than EUR 5 million. These companies should be independent (subsidiaries and controlled companies are not eligible), their incorporation should not be the result of a branch split or merger from a previous company, and they should not have distributed profits. In addition, firms must have an objects clause explicitly related to technological innovation and should fulfil at least one of the following requirements: an R&D expenditure ratio higher than 15%; at least one-third or two-thirds of staff holding a PhD or a master’s degree, respectively; ownership of legal rights for a patent or a software.

These benefits are not automatically applied: start-ups must first apply for inclusion in a ‘special section’ of the Italian Business Registry (*Registro delle imprese*). The benefits are valid from the date of registration and for all the time during which they are registered in the ‘special section’. Registration is revoked as soon as the company no longer fulfils the legal requirements. This may be because the start-up no longer has a sufficient character of technological innovation, e.g. because its primary activity has changed, or because it ‘grows out’ of innovative start-up status — registration automatically lapses 5 years after the date of incorporation of the company or when its turnover exceeds EUR 5 million per year.

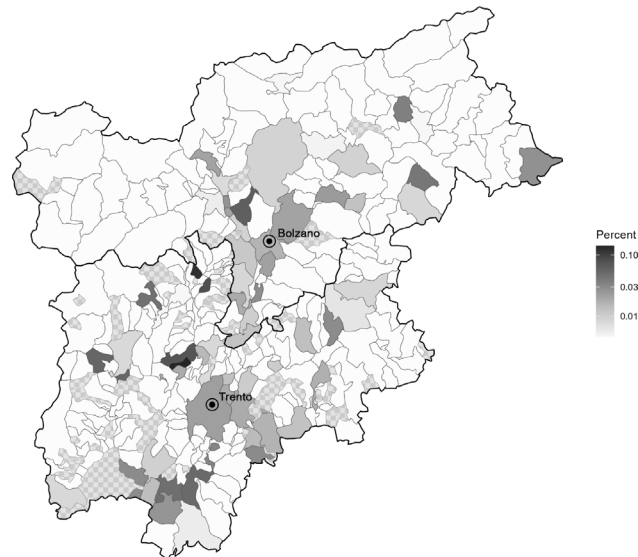
Self-selection is the distinctive feature of this regulatory framework: the benefits apply only to the eligible companies that voluntarily choose to register. This obviously implies that nascent and existing companies must be sufficiently informed about the policy to benefit from it, making its implementation vulnerable to knowledge gaps. A second feature of the Start-up Act is its ‘non-discriminatory’ approach: there are no economic sectors that are explicitly excluded, besides the requirement of carrying out a ‘technologically innovative’ business activity (Menon et al., 2018) — Moreover, there are no significant measures aimed at specific demographics, and most measures apply uniformly all over the country.

There are, however, several discernible patterns: start-ups are mostly concentrated in high-tech, low-capital sectors, such as digital services. When compared to the average Italian SME, start-ups display a higher concentration of younger entrepreneurs (OECD, 2020b) and are more common in the north of the country and, in general, in high-income areas (MISE, 2021). In Appendix B, Table 4, we provide a list of the ten most common sectors of economic activity, which together cover more than half of the total number of start-ups.

The policy has been fairly popular among small firms. Fig. 1(a) shows, as of 2019, the density of start-up firms relative to the total of firms in each region. While there is a clear start-up divide between the north and south of the country, the Trentino-Alto Adige region is among the top performers in terms of start-up creation. A less noticeable divide, however, appears to arise between the two provinces of Trento



(a) Registered Start-ups in Italy (%)



(b) Registered Start-ups in Trentino-Alto Adige/Südtirol (%)

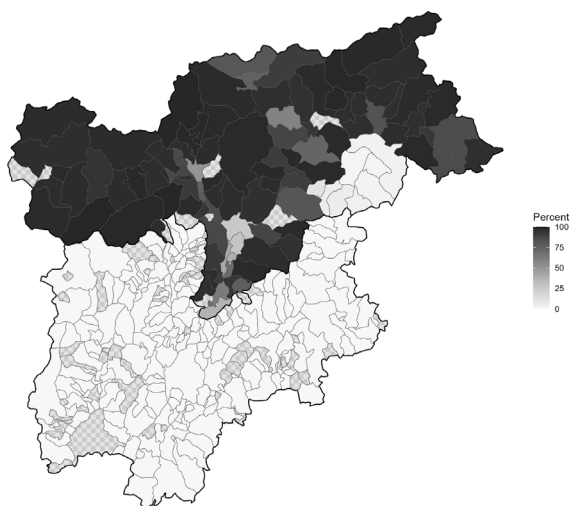
Fig. 1. Total registered start-ups as a percentage of firm population, among Italian regions 1(a) and the Trentino-Alto Adige/Südtirol provinces 1(b) (highlighted in a bold outline) and municipalities, as of 2019. Gradient in logarithmic scale. Authors’ calculations from AIDA and CAMCOM data. Municipalities with no firm data are marked with a checkered pattern.

and Bolzano. While the Trento province has one of the highest rates of innovative start-ups to firms, Bolzano’s rate is far lower yet still above the national average. In the next subsections, we discuss why that might be the case.

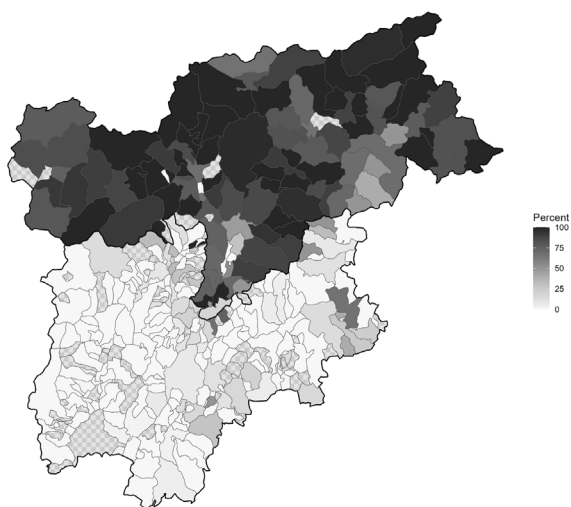
2.2. Ethnolinguistic divide in Trentino-Alto Adige

Differences in policy take-up seem to mirror an historically rooted cultural divide in the region. In fact, the region is composed of two neighbouring provinces: the provinces of Trento (‘Trentino’) and Bolzano (‘South Tyrol’). Trento is the southern province, while the official

⁶ Although, for the first couple of years, the age threshold was 4 years.



(a) Native German speakers in Trentino-Alto Adige/Südtirol (%)



(b) CEOs with German surnames in Trentino-Alto Adige/Südtirol (%)

Fig. 2. Language groups (2011) 2(a) and etymological origin of CEO surnames (2010–2019 average) 2(b) in Trentino-Alto Adige. Authors' calculations from census and AIDA data, using the surname classifier described in Appendix C. Municipalities with no firm data are marked with a checkered pattern.

name of the northern province, which shares its border with Austria, is Alto Adige/Südtirol, including both the Italian and German denomination.⁷

While Italian is the main language in Trentino, South Tyrol is the only Italian province where Italian is not the majority language. Each South Tyrol resident is required to be registered in one of the three language communities. Almost 70% of South Tyrol's population declares German as their mother tongue, with Italian native speakers making up around 25%. The remainder speaks Ladin, a distinct Romance language typical of the eastern valleys of the province. Linguistic shares in terms of proportion of German speakers are shown in Fig. 2(a). Fig. 2(b) also suggests that these same language patterns persist among entrepreneurs, as the distribution of German-named entrepreneurs closely follows the distribution of German speakers for each municipality.

⁷ South Tyrol is part of the historical region of Tyrol, which includes the Austrian territories of North and East Tyrol.

Trentino and South Tyrol are similar in many respects. They have a similar size in terms of residents⁸ and firm population,⁹ and are equally rural, with a large proportion of the population living outside the few mid-sized urban centres.¹⁰ They are among the richest provinces in the country, with South Tyrol ranking second and Trentino sixth in terms of gross domestic product across 100 Italian provinces,¹¹ and with a similar average household income¹² and economic performance (Banca d'Italia, 2021).

Nonetheless, there is a clear divergence between the two provinces in terms of registration to the Start-up Act's benefits scheme. Fig. 1(b) shows the density of registered firms relative to the total firm population in each municipality of the region. As the figure shows, the southern half, Trentino, is the area in Italy where the Start-up Act is most popular: by 2019, around 7.5% of all limited companies less than 5 years old registered as start-ups, by far the highest ratio in the country. Conversely, start-up density is much lower in South Tyrol: only 3.4% of all young limited companies registered as start-ups. While this statistic is in line with the national average, it is lower than most other affluent areas surrounding it. The lack of start-up firms on the west border of the region allows us to rule out spill-over effects from the neighbouring Lombardy, from which the Trentino province might benefit.

In addition, the start-up composition of the two provinces also differs in terms of geographical distribution. While start-ups are quite evenly distributed across the territory of Trentino,¹³ in South Tyrol, they are mainly concentrated in and around the provincial capital, Bolzano. While only 20% of the provincial population lives in the city, over two-thirds of its innovative start-ups are based there. These differences in registration rates are not transitory. As shown in Fig. 3, these trends have persisted over time, as the share of currently registered firms remains larger. The registration rate of firms with German-named CEOs has increased over time, but so has the rate of registration for firms with CEOs of Italian etymological ancestry.

Many of these divergences could be explained by differences in political priorities and industrial structure between provinces. Compared to Trentino, the industrial structure of South Tyrol has typically been more reliant on alpine tourism, construction and exports to Austria and Germany (Banca d'Italia, 2021). From an administrative perspective, both provinces enjoy the status of *provincia autonoma*, which entrusts them with a large degree of legislative, fiscal and budgetary autonomy compared to other Italian local authorities. There are differences in the level of support given to start-ups (in general) between the two provinces, with the Trento province having invested far more than Bolzano (Banca d'Italia, 2021). This aligns with the experience in the rest of the national territory: right after the implementation of the Start-up Act, heterogeneous programs supporting start-ups have been promoted at the regional level (Albanese et al., 2019).

Alternatively, these differences might be explained by differences in the formal and informal institutional contexts affecting the market economy of the two provinces, as evinced by popular views on

⁸ 545,000 inhabitants in Trentino; 535,000 in South Tyrol, according to the 2011 census.

⁹ The average difference between the two provinces in terms of the number of firms filing yearly accounts in each province was 433 between 2010 and 2019.

¹⁰ Istat (2015): Principali dimensioni geostatistiche e grado di urbanizzazione del Paese. Available at: <https://www.istat.it/it/archivio/137001> (last accessed 27 January 2022).

¹¹ Data is from 2018. Source: Eurostat, Gross domestic product (GDP) at current market prices by NUTS 3 regions.

¹² In the period 2009–2018, the average household income in South Tyrol and Trentino was EUR 38,000 and EUR 34,000 respectively. Source: ISTAT elaboration based on EU-SILC, Household average income. Available at <http://dati.istat.it/Index.aspx?QueryId=22919#> (Last accessed: 21 January 2022).

¹³ Reflecting local population patterns, the ecosystem is polycentric, with start-up agglomerations also in semi-rural areas (OECD, 2020b).

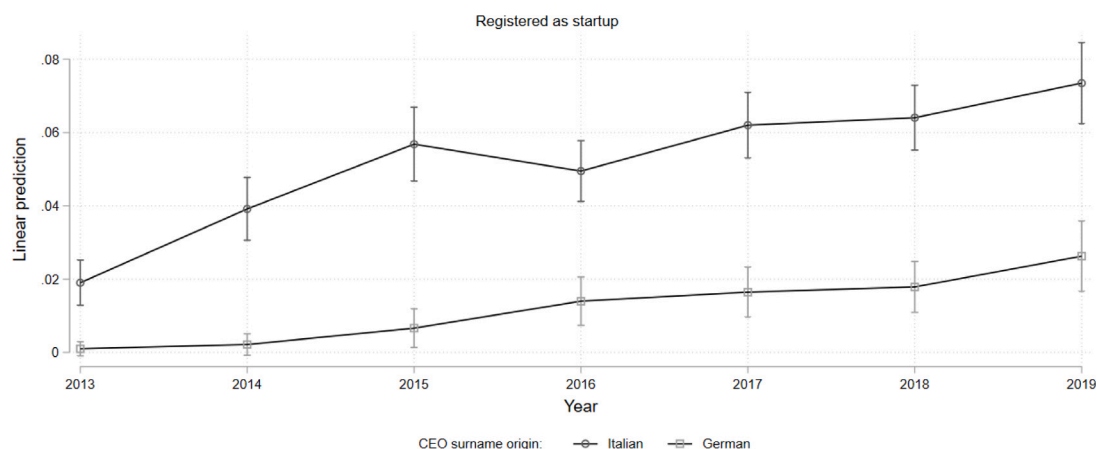


Fig. 3. Etymological origin of CEO surnames in Trentino-Alto Adige and probability of registration after the enactment of the Start-up Act. Linear prediction, using only year and surname origin (and their interaction) as predictors. Standard errors clustered by firm.

varieties of capitalism (Hall and Soskice, 2001) and institutional matrices (North, 1990). Evidence from the related literature (Baumol, 1990; Li and Zahra, 2012; Williams and Vorley, 2015; Eesley et al., 2018; Webb et al., 2020) suggests that these institutional factors might affect entrepreneurship dynamics and the propensity to innovate, and these institutional factors might be correlated with the language distribution.¹⁴

It could be tempting to attribute all differences in policy implementation solely to these territorial factors of an institutional and industrial nature. Yet, some facts suggest that some of these differences might go deeper than that, and might occur at the level of the firm itself. We discuss these in the next subsection.

2.3. Network embeddedness and information asymmetries

We propose that linguistic cleavages in the region have acted as an additional barrier to policy adoption, by giving rise to information asymmetries affecting the opportunity cost of registration between the two major ethnolinguistic groups.

In fact, it is clear that differences in registration rates have arisen not only within each province but also within each municipality. As it becomes apparent from Figs. 1(a) and 2(a), municipalities in South Tyrol with the highest density of Italian speakers are also more likely to feature registered start-ups. Start-up density correlates with the proportion of Italian speakers at the municipal level, even after controlling for population and urbanisation levels. These stylised facts, already noted by a 2020 OECD study (OECD, 2020a), cannot be explained solely by differences in provincial policies or industrial composition. As institutional and industrial factors are fixed for the population of entrepreneurs established in a given municipality, this variation in policy adoption must be explained with some other factors correlated with language.

Several factors might contribute to the emergence of these asymmetries. *Prima facie*, these asymmetries could be attributed to administrative barriers, such as limited availability of official documentation in German.¹⁵ We do, however, believe that these kinds of barriers have had a very limited impact. Quality and efficiency of the public administration in Trentino-Alto Adige are well above national standards (see

¹⁴ It might not be unreasonable to assume that the German-language areas feature a type of coordinated market economy that characterises German-speaking countries, rather than the mixed market economy that is typically attributed to Italian-speaking regions.

¹⁵ Available at: <https://www.camcom.bz.it/de/dienstleistungen/handelsregister/dienste-des-handelsregisters/innovative-start-und-kmu> (Last accessed: 21 January 2022).

e.g. Nifo and Vecchione, 2014). Not unlike any other law, the Start-up Act was translated into German in the Official Regional Bulletin, in accordance with national law (*Decreto del Presidente della Repubblica n. 574, July 15th 1988*). Furthermore, the local Chamber of Commerce has been offering German-language support to the firms seeking it. The voluntary nature of the Start-up Act reminds us that firms seeking this support need to be interested in the policy to begin with. This consideration makes us turn to a second, more nuanced hypothesis.

Network embeddedness (Granovetter, 1973) refers to the situation in which social ties between economic agents affect the sharing of resources and information within their network, along with the economic behaviour of the agents themselves. There is evidence for network embeddedness to affect the creation and absorption of new technologies (Gilsing et al., 2008) and that embeddedness in regional networks can have both positive (Cooke et al., 2005) and negative effects (Andersen, 2013; Gallo and Plunket, 2020) in terms of performance and knowledge diffusion, especially for SMEs.

We reckon that firms with Italian-speaking CEOs or Boards might innately benefit from a higher degree of embeddedness in the national network of firms than firms with German-speaking administrators. However, if this were solely the case, we would expect German-speaking administrators and their firms to be more open to new opportunities from the 'centre', in accordance with Granovetter's 1973 'strength of weak ties' theory. On the contrary, German-speaking administrators might also be too embedded in their own network (which embraces agents from Austria and Germany) to interact with relevant actors from other Italian regions.

There is evidence from surveys and experiments in support of the idea that, in Trentino-Alto Adige, language assignment gives rise to compartmentalised linguistic networks: German speakers tend to interact more with other German speakers, and so do Italian speakers (Vettori et al., 2012; Vettori and Abel, 2017). On a related note, Angerer et al. (2016) find that cooperative behaviour between the language groups tends to decrease with age, further insulating the two communities. Ethnolinguistic proximity is known to influence several other aspects of the socio-economic life of the region, such as tourist flows (Accetturo et al., 2019). Furthermore, loan applications in the regions appear to be influenced by cultural proximity on the credit demand side, with firms more likely to seek credit from institutions with a similar ethnolinguistic background, and this effect also appears to be increasingly stronger for small and young firms (Accetturo et al., 2023).

The idea that ethnic networks, in general, can also affect the flow of information is not new and holds an empirical basis (Larson and Lewis, 2017). In Trentino-Alto Adige, ethnolinguistic networks have already been found to affect the spread of online misinformation (Cantarella

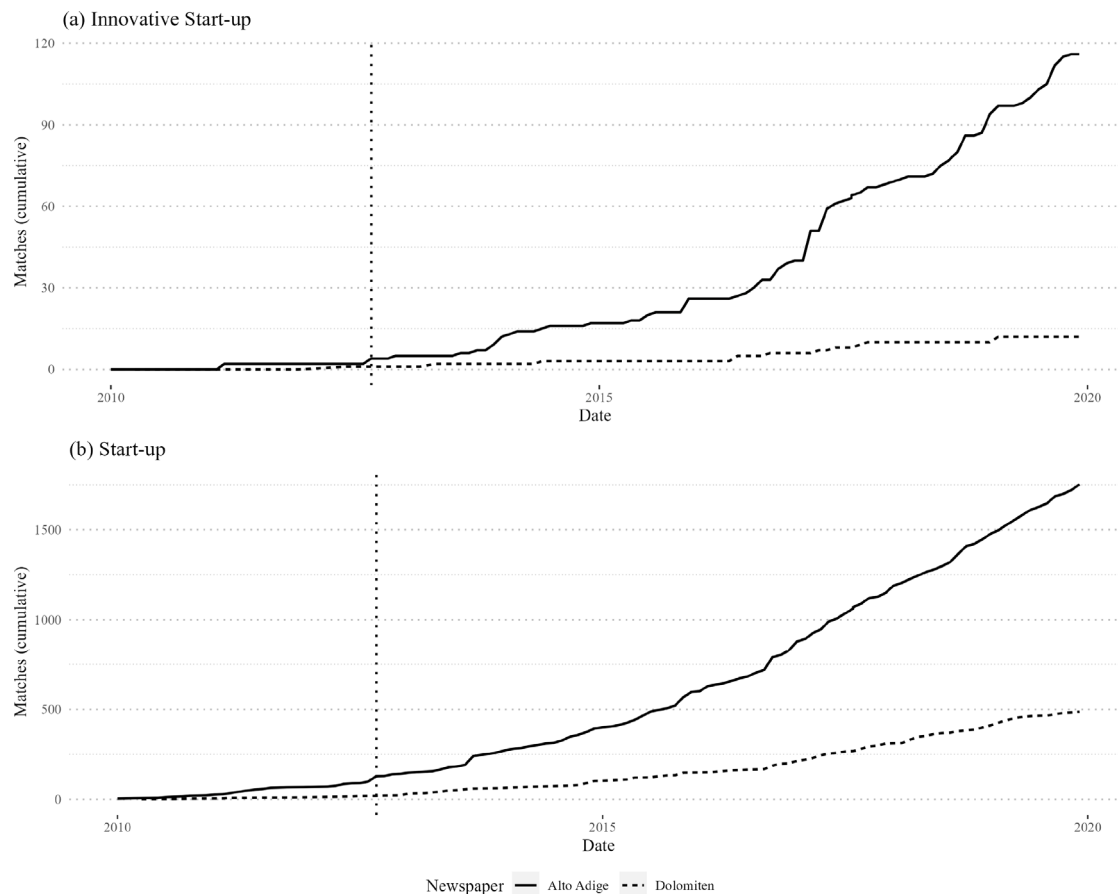


Fig. 4. Monthly cumulative mentions (2010–2019) of ‘Innovative Start-up’ (4a) and ‘Start-up’ (4b) in the Italian-language *Alto Adige* and the German-Language *Dolomiten* newspapers. Matches include ‘startup’, ‘start-up’, and ‘start up’ variations, along with plural and composite words. The dotted vertical indicates the introduction of the policy in October 2012.

et al., 2023) in the context of the 2018 national elections. The idea that the same process could affect the spread of information in other contexts is not unreasonable.

The role of network embeddedness for knowledge diffusion in the Start-up Act is substantiated by the fact that network agents themselves appear to be the first and principal source of information about the policy. In a survey from Istat/MISE (2018), 67.4% of the interviewed registered innovative start-ups revealed that tax accountants/financial advisors (figures known in Italy as *commercialisti*, ‘accountants’ henceforth) were a source of information about the Start-up Act. Among all other sources, accountants were ranked the highest, followed by online media, shared by 41.8% of interviewees. Institutional sources are ranked much lower, with only 25% of registered firms having received information about the Start-up Act from the Chambers of Commerce.

There are several ways in which ethnolinguistic network embeddedness can affect the diffusion of information on the Start-Up Act. These could range from differences in entrepreneurial ecosystems to institutional differences, to simply disinterest or mistrust towards the central government. Language proficiency might also play a role,¹⁶ but in high-tech, high-skill industries, language barriers should be easily overcome. What matters is that, no matter the source, information circulates differently among equally eligible firms, and that the attached policy benefits are, on paper, equally palatable to each

¹⁶ Although education is bilingual (both Italian and German are compulsory subjects for members of both language groups from age six), effective bilingualism is not widespread (see Vettori et al., 2021; Abel et al., 2012; Vettori, 2016; Coia et al., 2012).

group.¹⁷ Idiosyncratic variations in the *perception* of these benefits are still reduced to information asymmetries.

The greatest challenge, however, arises from disentangling linguistic effects from territorial and provincial differences in policies and industrial structure, which might make specific areas more or less susceptible to the production of eligible firms. These differences are still allowed to affect the supply of information of the Start-Up Act, but the demand for information should be kept fixed by ensuring that the firms under study are equally eligible. In other words, the non-registration of a German-language firm can be attributed to industrial differences only as long as these affected the information accessible to this firm and not because the firm never had the potential to become an innovative start-up to begin with. A central objective of this paper is to test whether asymmetries in registration are robust to regional sources of heterogeneity correlated with language.

3. Data

Our research methodology combines various datasets and surveys to analyse policy coverage, awareness, and take-up separately. The policy take-up analysis relies on business data from Bureau van Dijk’s Aida database, with a focus on eligible firms located in Trentino-Alto Adige in the 2010–2019 window. Additionally, open data from the Italian Business Registry is used to reconstruct the history of companies

¹⁷ Which is obviously the case since the benefits deriving from registration to the policy are *automatic* and apply in areas which are beyond the scope of regional autonomy — e.g. corporate law, business taxation, and access to private finance.

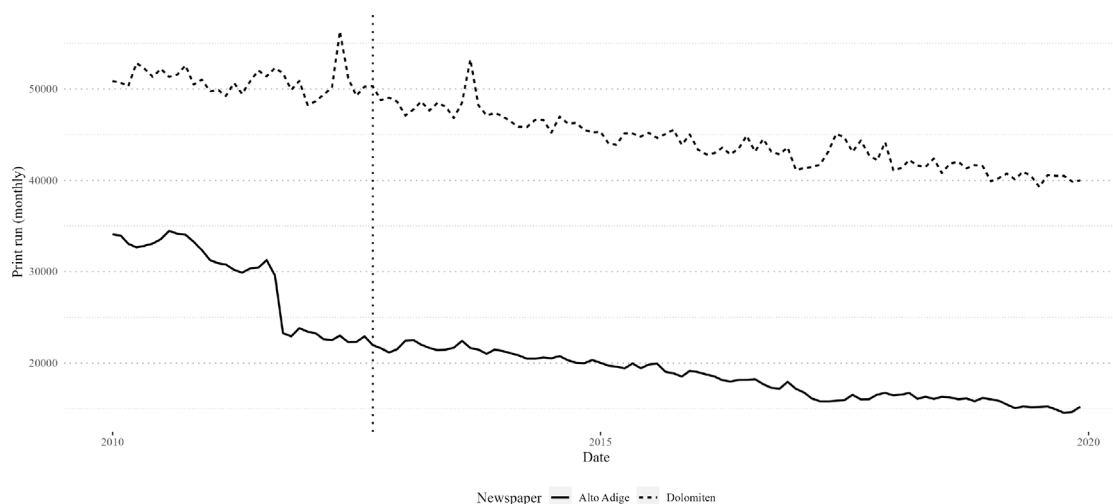


Fig. 5. Monthly print runs (2010–2019) of *Alto Adige* and *Dolomiten*. Paper and digital copies.

registered as innovative start-ups, and text analysis methods are used to detect the etymological origins of the surnames of CEOs and board members, as detailed in Appendix C. The sample was carefully curated to select eligible firms only, ensuring that only firms that could have realistically acquired the innovative start-up status are included, as detailed in Appendix D. We also match our data with firm-level information on access to provincial and EU funding, using publicly available data from *Open Coesione* on beneficiaries of regional programs co-financed by *ERDF* and data on provincial subsidies for start-up creation from *Accetturo* (2022).

The study also collects data on policy awareness through a survey conducted in 2023 among residents of Trentino-Alto Adige, covering knowledge of various national policies, including the Start-up Act. Furthermore, textual data from local newspapers and online sources specialising in business and entrepreneurship are analysed to study the supply of information on the Start-up Act during the 2010–2019 window.

All these data sources are described in detail in Appendix A.

4. Policy coverage in news outlets

There are a few indirect ways to test if the spread of information on the Start-up Act has been affected by linguistic networks, and coverage of the policy in local newspapers can provide such a test. By analysing German and Italian-language newspapers in the area, we investigate how often the reform was covered in the press. To find support for the information asymmetry hypothesis, we would expect fewer mentions of the policy in the German-language press.

Furthermore, news outlets are also an important *direct* source of information for firms about the Start-up Act. According to the *Istat/MISE* (2018) survey on innovative start-ups, 41.8 and 17.1% of firms have obtained information from the Start-up Act from online and in-print media sources, respectively. This analysis can then offer insights into the supply of information about the policy and not just into the presence of information asymmetries.

We analyse two pairs of generalist and business outlets. We have preferred to analyse newspapers rather than rely on Google Trends (which is a common tool for these types of tasks) for two reasons. First, Google searches are endogenous to the linguistic network and would capture the demand for information rather than the supply of information, which is what we are actually interested in. Second, Google Trends data do not offer a sufficient level of regional disaggregation in Trentino-Alto Adige and do not allow us to compare the trends between the two linguistic groups.

We compare each pair of newspapers by looking at differences in trends in the cumulative use of the terms ‘start-up’ and ‘innovative start-up’ in each Italian and German-language newspaper. ‘Start-up’ is a loan word for Italian and German, requiring no translation or searching for synonyms. The word is also used similarly in both the Italian and German languages: a quick analysis on Google Books Ngram Viewer reveals that, in 2008 (the last available year), the term ‘start-up’ and its variations appeared around 0.00006% times in recorded German-language texts and 0.00005% in Italian-language texts.¹⁸ These differences are basically insignificant and, hint that the direction of the bias would, if anything, favour the German language. For what concerns the term ‘innovative’, this same word is serendipitously used in both the official nomenclature in Italian (‘start-up innovative’) and its official German translation (‘innovative start-up’).¹⁹ Variations of the terms accounting for singular forms and composite words have nonetheless been included in the search.

Starting with the generalist newspapers, Fig. 4 compares the trends in coverage of the policy (and start-ups in general) across the two main newspapers in the mixed-language South Tyrol province: the Italian-language *Alto Adige* and the German-language *Dolomiten*. Both newspapers have a large print run for the region (historical trends are shown in Fig. 5), which mirrors the relative size of each linguistic group in the South Tyrol province. Between 2010 and 2019, *Alto Adige* prints amounted to around 31% of all print runs between *Dolomiten* and *Alto Adige*, a proportion close to the 25% circa Italian-speaking residents in the province.

The top Fig. 4a shows the cumulative sum of all mentions of ‘innovative start-ups’ (e.g. a nomenclature simultaneously used in both the bureaucratic and ‘PR’ parlance of the policy), while the bottom Fig. 4b shows the same trends for any mention of ‘start-up’. Starting from Fig. 4a, it appears that the policy has benefited from much more coverage by the Italian-language press since its introduction in late 2012. While it is possible that some ‘innovative start-up’ mentions might be unrelated to the policy, the nearly complete absence of mentions in the period preceding the policy also suggests that differences in trends are not to be attributed to community-specific noise.

Eventually, at the end of 2019, ‘innovative start-ups’ were mentioned 116 times in the Italian-language newspaper, against only 12

¹⁸ These figures can be retrieved at the following links: IT: <https://rb.gy/edkdd>; DE: <https://rb.gy/qomsz>, last accessed: 03/03/2023.

¹⁹ See, for example: <https://www.camcom.bz.it/de/dienstleistungen/handelsregister/dienste-des-handelsregisters/innovative-start-und-kmu>, last accessed: 03/03/2023.

mentions in the German-language one. Worthy of note is the fact that on the month the policy was introduced (October 2012), *Alto Adige* dedicated a full article on the policy (hence the two matches), while *Dolomiten* did not even cover this piece of news.

As many start-ups are involved in innovation of some form, it is unsurprising that the term ‘start-up’ alone is often used as a synonym for ‘innovative start-up’. Therefore, while it cannot be used as a counterfactual, it can be worthwhile checking for any growth in mentions of ‘start-ups’, as this will be correlated with the spread of information on the Start-up Act. Prior to the introduction of the policy, *Alto Adige* already used the term ‘start-up’ with slightly more frequency than *Dolomiten*. Yet, mentions remained minimal for both newspapers, with a difference of 110 cumulative mentions between the two newspapers. The difference grows as large as 1265 cumulative mentions at the end of 2019.

What is more interesting is the change in the rate of growth in cumulative mentions, indicating that since late 2012 a quantitative shift in the interest in ‘start-ups’ has taken place for the Italian-language newspaper. This is easily visualised by the change in the slope of the curve of mentions from October 2012 onwards, which becomes much steeper.

Most importantly, for *Alto Adige*, the shape of the curve of ‘start-up’ mentions closely resembles the shape of the ‘innovative start-up’ curve: this indicates that much of the growing news coverage on start-ups can be explained by the growing interest in the firms in the policy registry. This is far from the case for *Dolomiten*, for which the growing coverage of start-ups was not accompanied by an increase in coverage of innovative start-ups.

Notwithstanding the small pre-policy differences, there is then clear evidence of a differential change in coverage trends for both innovative start-ups and start-ups in general. Note, also, that not only is information about the Start-up Act shared less frequently in the German-language press, but new mentions of the policy also appear to be statistically uncorrelated with mentions in the Italian-language press. This evidence supports the idea that information flows differently between the two communities, suggesting that embeddedness in each linguistic network makes each community less permeable to the exchange of information with the other.

These findings apply to the two main generalist newspapers in the area. However, one could argue that CEOs do not necessarily consume news from local newspapers. We then focus on trends in coverage across specialised business outlets in Fig. 6. In this case, there is no regional Italian-language outlet targeted to entrepreneurs and tax accountants from Trentino-Alto Adige, but there is one in German. We then focus on the Bolzano-based, German-language *Südtiroler Wirtschaftszeitung* (SWZ), and the Milan-based Italian-language *Italia Oggi*, which we expect to be followed by Italian-speaking entrepreneurs all over the national territory. These newspapers are available online, and the search function can be used for free. We then recorded the title and date of publication of all articles matching our search queries through a web scraping algorithm.²⁰

Again, the analysis confirms our previous findings. The policy barely received any coverage in the German-language outlet, with no more than 33 articles having mentioned innovative start-ups by the end of 2019. Furthermore, the first mention of the policy occurred in July 2013, long after the introduction of the policy (the next mention was in 2014). The differences in coverage with *Italia Oggi* are striking. Also, while the indicator used in the figure is not comparable to the one used in Fig. 4, it is worth noting that even the generalist Italian-speaking local press (*Alto Adige*) has dedicated considerably more articles to the policy than SWZ did.

²⁰ Recording the exact number of mentions of the queried terms was, instead, not possible.

It could be argued that, given the local nature of SWZ, it is possible that there might be, in general, less interest in start-ups among German-speaking readers relative to the country-wide interest. This appears to be the case, as the coverage of start-ups in general (bottom figure) has been much larger for *Italia Oggi*. However, the difference between the trends in start-up coverage between the two outlets is not sufficiently large to justify the difference in policy coverage. On the day of the introduction of the policy, cumulative coverage of start-ups was two times larger on *Italia Oggi* than on SWZ. At the end of 2019, it was three times larger. Again, start-up coverage is endogenous to the coverage of innovative startups, especially when the two terms are used interchangeably, but this does not change the fact that, at the end of 2019, coverage of the policy was nearly twelve times larger in *Italia Oggi* than on SWZ.

The evidence presented strongly supports the idea that information asymmetries within the German-speaking network hamper the spread of information about the Start-up Act. However, the Italian-speaking press seems to be much more susceptible to the start-up discourse, even if these differences are not large enough to justify the near-complete disinterest in the policy among the German-speaking press. A possibility is that these differences might be driven by variations in the industrial tapestry of the province of South Tyrol, as German readers are much more likely to live in rural areas than Italian readers. The key question here relates to whether the lower coverage of start-ups, in general, is also symptomatic (if not causally related) to a lower probability for German-language entrepreneurs to create start-ups eligible for the policy, simply because these entrepreneurs are more likely to be based in rural areas. We will explore these questions in the next sections.

5. Policy awareness and ethnolinguistic diversity

5.1. Econometric model

To study whether differences in policy coverage have sedimented into differential awareness among the two language groups, we analyse survey data from the region, probing respondents’ policy knowledge on five different national policies in 2023, including the Start-up Act.²¹ This exercise allows us to evaluate whether ethnolinguistic assignment is also tied to differences in policy awareness and to understand which individual predictors may mediate, diminish or annul this assignment effect. In addition, the 2023 fieldwork allows us to understand if differences in awareness have persisted over time, far from the enactment of the policy, while focusing on a larger set of policies allows us to understand if the effect of assignment on knowledge of the Start-up Act is generalisable to other contexts, netting out policy-specific awareness effects.

In order to study the linguistic determinants of policy awareness, we need to separate the effects of idiosyncratic awareness from those of policy-specific awareness. To do so, we reshape the sample so that each of the five policies is treated as a different observation p for the same individual i , leading to a final sample of $N \times 5$ observations.

With this sample, we propose the following baseline multinomial logistic model:

$$Awareness_{imp} = \beta_p + Language_i \delta + Controls_i \zeta + Bilingual_i \gamma + e_{imp} \quad (5.1)$$

where $Awareness_{ip}$ is the vector of the three discrete outcomes related to self-reported awareness of the policy: ‘No knowledge’, ‘By name only’, and ‘Good knowledge’. $Language_i$, instead, is the reported native language of the respondent used for the completion of the survey,²²

²¹ These are: (i) a Citizens’ Income Scheme (“Reddito di cittadinanza - RdC”), (ii) Early retirement schemes (“Quota 100”, “Opzione donna”), (iii) Bonuses for young adults (“18app”), (iv) Incentives for the construction sector (“Superbonus”, “Bonus facciate”), and (v) The Start-up Act itself.

²² Surname information is, instead, not available.

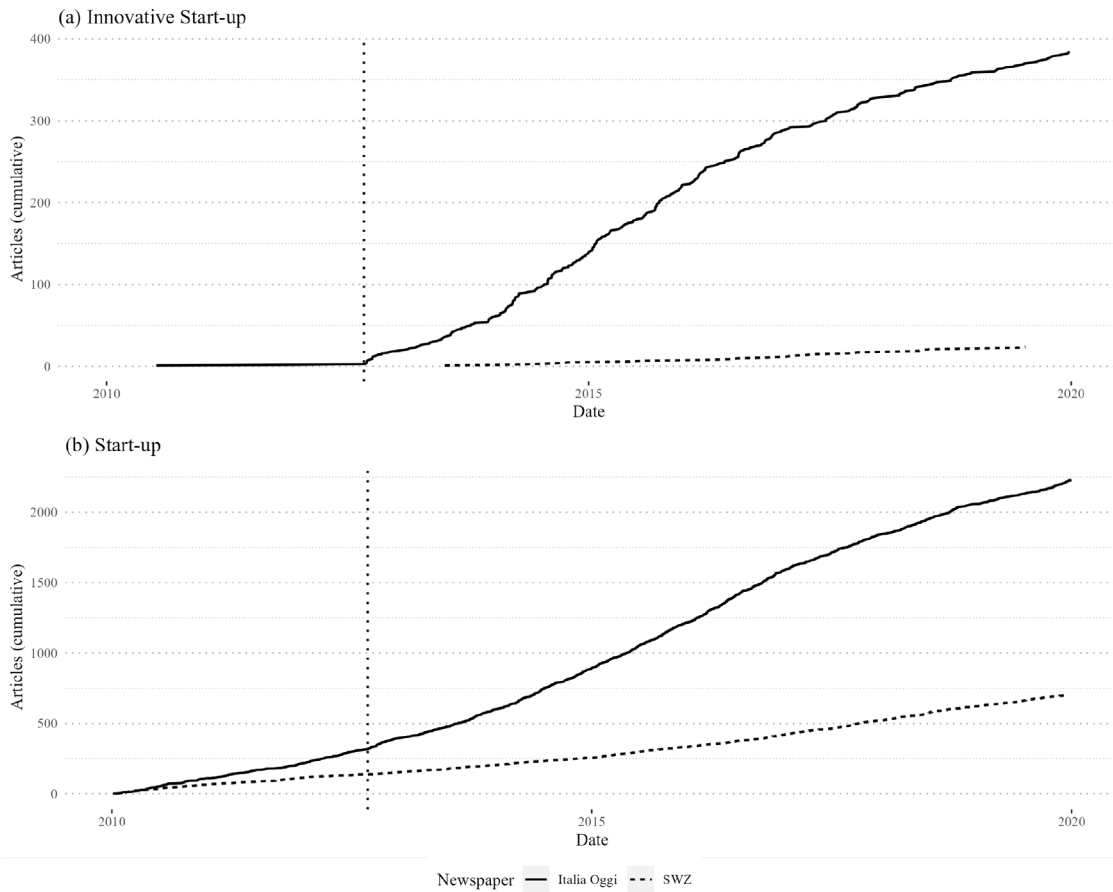


Fig. 6. Daily cumulative articles (2010–2019) mentioning ‘Innovative Start-up’ (6a) and ‘Start-up’ (6b) in the Italian-language *Italia Oggi* and the German-Language *Südtiroler Wirtschaftszeitung* online business newspapers. Matches include ‘startup’, ‘start-up’, and ‘start up’ variations, along with plural and composite words. The dotted vertical indicates the introduction of the policy in October 2012.

while $Bilingual_i$ is a dummy variable denoting whether the individual is also a native bilingual speaker. The vector $Controls_i$ captures basic demographic information (gender, age, marital status, and education).

β_p is the vector of policy-specific effects on awareness. Each intercept of this vector will reveal if a policy is more or less obscure to the general population.

This baseline model ignores local sources of variation in policy awareness, which might be correlated with language. To account for these, we update the model as follows:

$$Awareness_{imp} = \alpha_m + \beta_p + Language_i \delta + Bilingual_i \gamma + Controls_i' \zeta + Embeddedness_i' \theta + e_{imp} \quad (5.2)$$

The new controls are meant to capture four dimensions of embeddedness with the opposite language group: (i) the second-language skills, (ii) the place of residence, (iii) the physical presence of members of the other language group, and (iv) the daily use of the non-native language.

The place of residence dimension is captured by α_m , the municipality fixed effect, through the postal code of the resident, while the vector $Embeddedness_i$ captures the remaining dimensions, by controlling for the proficiency in the second language, the share of Italian/German-speaker among the three closest neighbours to the household’s main residence, and the frequency of use of the second language with complete strangers during the last week.²³ These controls are introduced in a stepwise fashion to appreciate which of these dimensions can better explain information asymmetries.

²³ This is captured on a 5-point scale, from (i) never, (ii) once or twice, (iii) three of four times, (iv) five or six, to (v) every day.

5.2. Results

Descriptive results from the survey are reported in Fig. 7, showing the distribution of knowledge of each of the five policies as a percentage of the population of each ethnolinguistic group.

While there is substantial variation in awareness between policies, the figure suggests that, in almost any case, German-speakers tend to self-report as less knowledgeable about national policies, and these tendencies persist among all policies under analysis. However, these results, alone, could be deceiving, as language might mediate other characteristics which might ultimately explain differences in policy awareness.

Multinomial results for policy awareness are then reported in Table 1, holding ‘Good’ policy knowledge as the baseline outcome, and the citizen’s income scheme ‘Reddito di Cittadinanza’ (RdC) as the baseline outcome. Beginning with the baseline model in specification (1), we immediately find a strong and statistically significant effect of language group assignment on policy awareness. The multinomial log-odds for by-name-only knowledge relative to good knowledge would be expected to decrease by 0.54 units, holding everything else constant. Similarly, the log-odds for no knowledge suggest a decrease in awareness by 0.81 units, relative to good knowledge.

We introduce in model (2) the first of the embeddedness controls, starting with individual skills in the second language. While the effect is strong and statistically significant, pointing at a negative effect of language skills on policy awareness, it is not sufficient to affect the magnitude and sign of the main language coefficient. In any case, it could be argued that second-language skills are not an exhaustive measure of embeddedness, as they capture the linguistic competence

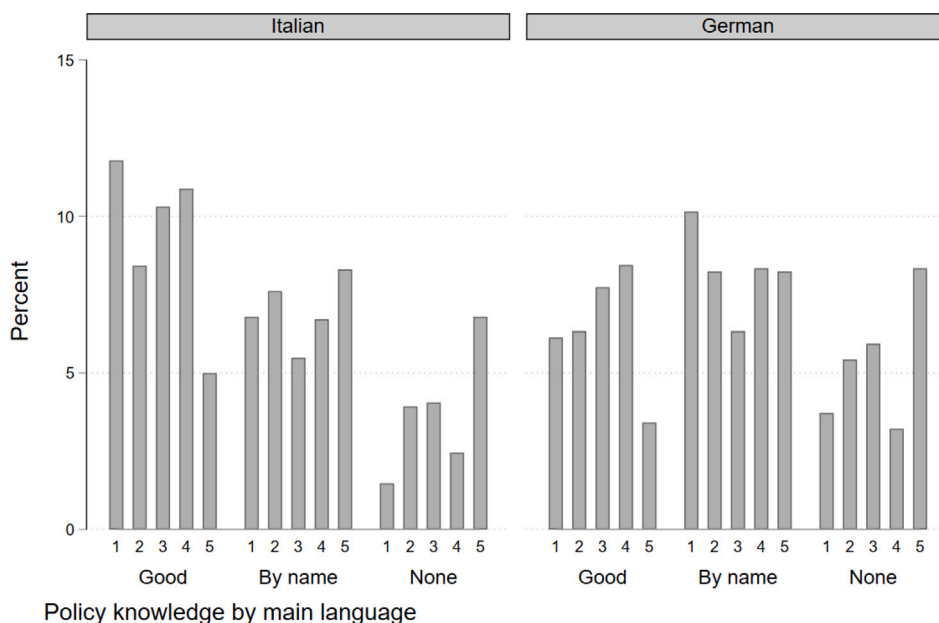


Fig. 7. Policy awareness by language group and policy type. Policies: (1) RdC, (2) Quota100, (3) 18App, (4) Superbonus, (5) Start-up Act.

effect on awareness for *both* groups, suggesting that lower language skills in German for Italian-speaking respondents are also connected with poor policy knowledge.

We then look at other determinants of ethnolinguistic embeddedness. In model (3), we introduce the location f.e. using the postal code of the municipality of residence. While their inclusion contributes to a substantial increase in the model fit, these intercepts do not seem to have a strong mediating power on the main language effect on policy awareness, which remains nearly identical to the estimates from model (1). Model (4) adds controls for the linguistic composition of the household's neighbours. However, their effect is far from statistically significant, presumably because these are highly correlated with the location f.e.

These results do not necessarily tell us whether respondents' language skills and linguistic context are connected with higher usage of the second language. In column (5), we add our final set of controls by including fixed intercepts for the use of the second language in the respondent's daily lives. Namely, we control for how often the respondent has made use of the second language over the last week, conditional on their main language.

The inclusion of the language use controls significantly affects our estimates by rendering the effect of native language statistically null. Furthermore, the fact that we are already controlling for all other embeddedness controls suggests that this effect cannot be attributed to language skills and location factors that already affect the overall frequency of interaction with individuals from the two language groups. These results are extremely important because they suggest that the connections between individuals, mediated by assignment to a native language, are the main predictors of policy awareness, net of language skills and location.

There seems to be considerable within-policy variation in awareness. Compared to the Citizen's income ('Reddito di Cittadinanza - RdC') baseline, the Start-up Act features a much higher likelihood of being unknown to the general population. This is not unexpected, given that the policy targets a specific population of entrepreneurs that is very different from the general public.

However, we might wonder if our results hold for the Start-up Act specifically, and we might wish to focus on the heterogeneous effects of linguistic assignment on awareness of the Start-up Act specifically. We do so in Table 9 in Appendix G. Dropping the other policies leads to a smaller sample and larger variability in our estimates. Yet, the

persistence of the linguistic effect until the final set of language use controls is introduced (which even inverts the sign of the coefficient, albeit only at the 5% level of significance) supports all our findings so far. Furthermore, the fact that differences in policy adoption (until language use is accounted for) have endured in 2023 for a policy implemented exactly a decade prior suggests that these informational asymmetries are persistent and that, in fact, no catching-up has taken place over the last ten years.

In conclusion, our results suggest that ethnolinguistic assignment unambiguously affects awareness and knowledge of any national policy, and the Start-up Act specifically. The fact that the language effect persists even after controlling for each type of policy and within postal codes, together with the fact that respondents are not entrepreneurs, suggests that this awareness mechanism goes beyond differences in local priorities and institutions.

Only after controlling for interactions with members of the opposite language group is the ethnolinguistic effect made null. However, as native-language assignment predates language use, these results should not suggest that the ethnolinguistic effect is spurious, but rather that ethnolinguistic assignment affects policy awareness through the major channel of network embeddedness. What our results show is that linguistic assignment coalesces into information asymmetries, but only when they affect interactions between members of different communities.

6. Policy take-up and ethnolinguistic diversity

6.1. Econometric model

In this section, we study the effect of language assignment on registration examining whether firms with German-speaking administrators are factually less likely to adopt the policy, and showing under which conditions language can fully be treated as exogenous to a firm.

Using our Aida panel, we focus on eligible firms aged less than five years (four until 2015, when the age threshold was extended), treating registration status as a time-variant factor, which 'switches on' once it acquires the start-up status, and it stays this way until it is removed from the sample due to age ineligibility.²⁴ We estimate the effects of

²⁴ For example, for a hypothetical firm that registers in 2015, the treatment status is 0 until the year 2015, after which the status switches to 1 until the firm ceases to qualify as a start-up because of the age threshold.

Table 1
Main language and policy awareness.

Variables	(1)		(2)		(3)		(4)		(5)	
	MLOGIT		MLOGIT		MLOGIT		MLOGIT		MLOGIT	
	By name	None	By name	None	By name	None	By name	None	By name	None
Main language: German	0.543*** (0.126)	0.808*** (0.155)	0.773*** (0.138)	1.184*** (0.176)	0.533*** (0.155)	0.831*** (0.199)	0.496*** (0.170)	0.831*** (0.231)	0.182 (0.261)	0.045 (0.373)
Policy: Quota100	0.279*** (0.094)	1.016*** (0.129)	0.286*** (0.095)	1.027*** (0.131)	0.297*** (0.097)	1.054*** (0.134)	0.302*** (0.098)	1.088*** (0.139)	0.310*** (0.099)	1.092*** (0.140)
Policy: 18App	-0.238** (0.114)	0.856*** (0.146)	-0.234** (0.114)	0.863*** (0.148)	-0.236** (0.117)	0.885*** (0.152)	-0.231* (0.118)	0.910*** (0.157)	-0.224* (0.119)	0.913*** (0.159)
Policy: SuperBonus	-0.080 (0.097)	0.234 (0.153)	-0.079 (0.097)	0.235 (0.155)	-0.080 (0.100)	0.239 (0.158)	-0.092 (0.101)	0.259 (0.164)	-0.092 (0.102)	0.261 (0.165)
Policy: Start-up Act	0.909*** (0.119)	2.117*** (0.152)	0.925*** (0.119)	2.144*** (0.153)	0.957*** (0.122)	2.210*** (0.159)	0.983*** (0.123)	2.271*** (0.162)	0.992*** (0.123)	2.295*** (0.163)
Bilingual	0.080 (0.164)	-0.038 (0.220)	0.287 (0.176)	0.268 (0.236)	0.255 (0.181)	0.241 (0.235)	0.256 (0.185)	0.216 (0.243)	0.293 (0.192)	0.283 (0.248)
Age	0.002 (0.004)	-0.013** (0.006)	0.001 (0.004)	-0.015** (0.007)	0.001 (0.004)	-0.013** (0.007)	0.002 (0.004)	-0.012* (0.007)	0.002 (0.004)	-0.012* (0.007)
Gender: female	0.191* (0.114)	0.151 (0.149)	0.210* (0.113)	0.178 (0.149)	0.220* (0.116)	0.209 (0.157)	0.250** (0.117)	0.274* (0.157)	0.254** (0.118)	0.228 (0.159)
2nd-language proficiency			-0.122*** (0.035)	-0.201*** (0.045)	-0.132*** (0.038)	-0.240*** (0.049)	-0.126*** (0.038)	-0.231*** (0.049)	-0.090** (0.043)	-0.149*** (0.055)
Neighbours: % Italian-speaking							-0.275 (0.197)	-0.313 (0.313)	-0.258 (0.199)	-0.234 (0.321)
Neighbours: % German-speaking							-0.166 (0.215)	-0.299 (0.309)	-0.143 (0.226)	-0.271 (0.326)
Observations	3439	3439	3439	3439	3439	3439	3399	3399	3394	3394
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Language use controls	No	No	No	No	No	No	No	No	Yes	Yes
Pseudo R-squared	0.0575	0.0575	0.0641	0.0641	0.0879	0.0879	0.0897	0.0897	0.0954	0.0954

Notes: SE clustered by ID in parentheses. Base outcome: “Good knowledge”; Base policy: “Policy 1: RdC”. Other controls: see Section 5.1.
*p<.05; **p<.01; ***p<.001.

language assignment at the same time point as the outcome, without any lags or leads (with the exception of some lagged control variables, see below for further details).

As the actual CEO (and firm) language is unobserved,²⁵ our econometric strategy studies the reduced form effect of ethnolinguistic assignment on policy adoption over the years. We exploit the etymological root of surnames as a proxy for language assignment, which captures the intention to treat better than the language spoken by a CEO: while language proficiency is endogenous as it can be acquired throughout life, surnames are assigned at birth and therefore do not depend on CEO’s personal characteristics or choices. This accounts for the possibility of a CEO being bilingual or fluent in a non-native language, which might correlate with other factors such as skill, income or education.²⁶

Our main concerns relate to whether conditional independence of surnames is obtained: in short, we need variation in language assignment to originate from the firm only, and not from regional factors. In fact the main issue related to the surname spatial distribution is it can be correlated with local factors affecting economic policies and industrial structure, which could undermine the validity of our results. This might be the case, for example, of provincial and municipal variation in public funding, differences in entrepreneurial ecosystems, and any other source of variation above the firm level that simultaneously predicts language assignment and affects the likelihood of registration. The advantage is that these factors are only space–time-variant, so they are easy to control: municipalities cannot move in space or time, and industry classes are also fixed. We develop three estimation strategies, each testing the robustness of the connection between ethnolinguistic diversity and policy adoption under different assumptions.

²⁵ AIDA does not offer information on the native language of CEOs or board members.

²⁶ A CEO might belong to the German linguistic group and have an Italian surname or vice versa. This is not an issue for the estimation of the surname effect as long as the connection between surnames and language stays strong enough: as shown in Fig. 2, this appears to be the case.

To appreciate how these location factors can affect our estimates, we start with the baseline model:

$$Registered_{it} = \beta_i + Surname_{it}\delta + e_{it} \tag{6.1}$$

and then test additional specifications in which we try to absorb these sources of variation in language assignment that are external to the firm. *Registered* is a binary treatment indicator that takes the value of one if a firm has registered as an innovative start-up, and zero otherwise. The dummy variable *Surname* captures the effect of interest. This variable equals one if the surname of the CEO of firm *i* has German (Germanic) origins and zero if it has Italian (Romance) origins. The variable is also switched with the *BoDSurname* variable, denoting the share of German-named members in the firm’s board of directors, in all our main specifications. β_i are time intercepts.

Appendix E discusses in detail the main conditions under which our identification strategy holds, while Appendix G offers additional robustness checks by introducing the territorial-industry fixed effects in a stepwise manner and offering checks for firms’ access to other funding sources. Overall, our tests support the idea that net of municipal and provincial variation, surnames predict ethnolinguistic assignment yet fail to explain differences in performance among firms.

As surnames are rendered exogenous, these cultural asymmetries might be used to study the effectiveness of said policies. We explore this possibility in Appendix F, in which we use the model from this section as the first stage of a two-stage model to evaluate the impact of the policy in the region. While this exercise is valid for the region under analysis, its findings could not necessarily be externally valid, i.e. apply to the full country. However, they are informative about the effectiveness of the policy in Trentino Alto Adige and contribute to the debate on the causal impact of the Start-Up Act on firms.

6.1.1. Saturated fixed effects

In our first specification, we employ a saturated fixed effects strategy that holds municipality, industry, year and their full set of interactions constant. As a result, this model ensures that specific policies

that target specific industries, in specific municipalities, and in specific years cannot possibly drive our results.

We then estimate the following reduced-form specification over the full sample of eligible firms i in the region between 2013 and 2019 for municipality m , industry n , and year t :

$$\text{Registered}_{imnt} = \alpha_m + \beta_t + \gamma_n + \kappa_{mt} + \lambda_{nt} + \mu_{mn} + \nu_{mnt} + \text{Surname}_{it}\delta + \text{Controls}'_{it}\zeta + \text{Financials}'_{it-1}\eta + e_{imnt} \quad (6.2)$$

We include the vector *Controls* to control for firm-level characteristics. These include board size and CEO characteristics (age, gender, foreign nationality, foreign nationality \times age). These controls are all interacted with dummies for the years of activity of a firm to account for heterogeneity related to firm age. In the *Financials* $_{t-1}$ vector, we include the logarithm of firm time-variant factors (lagged assets, bank debt and number of employees, and with dummies for each of these variables being zero) at $t - 1$, along with three lagged proxies of the optional access criteria. These are R&D expenses, wages, and patents at $t - 1$, which are also fully interacted to allow for all possible combinations of access criteria.²⁷ The saturated fixed effects allow us to control for a number of unobserved factors. α_m will capture unobserved time-invariant municipality effects, while year effects are captured by β_t . We control for unobserved time-variant heterogeneity across municipalities by interacting the two in κ_{mt} . Industry effects are captured by the parameter γ_n , along with their interactions with time λ_{nt} and municipality μ_{mn} . Finally, the interaction cells between these three variables are also included in the intercept vector ν_{mnt} . Province-specific factors, along with time-variant policies, are collinear with municipal factors and are then fully captured as well.

Our specification belongs to the family of staggered DiD. As such, it relies on the assumption that the treatment effect is stable (de Chaisemartin and D'Haultfoeuille, 2018) and homogeneous, so as not to incur the negative weights problem. Eq. (6.2) allows for heterogeneous assignment effects, as long as they are generated at a level higher or equal to the municipality or industry. In short, it relies on the assumption that the firm-level linguistic assignment effect is fixed after controlling for higher-level sources of language variation in space and time. However, we can relax this assumption and, following from Wooldridge (2021), replace the language effect coefficient with the vector of time-varying coefficients δ_t , yielding the equation:

$$\text{Registered}_{imnt} = \alpha_m + \beta_t + \gamma_n + \kappa_{mt} + \lambda_{nt} + \mu_{mn} + \nu_{mnt} + \text{Surname}_{it}\delta_t + \text{Controls}'_{it}\zeta + \text{Financials}'_{it-1}\eta + e_{imnt} \quad (6.3)$$

6.1.2. Difference in differences: Saturated fixed effects

The previous specification did not explicitly absorb differences between language groups. To offer an alternative model in which these differences are explicitly estimated, we develop a difference in differences (DiD, henceforth) strategy to hold differences among firms belonging to different linguistic groups as fixed.

Including available pre-treatment years, we estimate the following equation over all eligible firms between 2011 and 2019:

$$\begin{aligned} \text{Registered}_{imnt} = & \alpha_m + \beta_t + \gamma_n + \kappa_{mt} + \lambda_{nt} + \mu_{mn} + \nu_{mnt} \\ & + \text{Surname}_{it}\theta + \text{Surname}_{it} * \text{Post}_t\delta + \\ & \text{Surname}_{it} * \text{Trends}'_{mt}\phi + \text{Trends}'_{mt}\psi \\ & + \text{Controls}'_{it}\zeta + \text{Financials}'_{it-1}\eta + e_{imnt} \end{aligned} \quad (6.4)$$

This specification is similar to the previous one, except that the language group effect on registration is now captured by the interaction

between *Surname* and the post-treatment dummy *Post*, denoting post-treatment years. We add to the *Trends* vector observable time-varying municipality controls. These controls also include the logarithm of the average of all financials in year t and municipality m for all firms except i . In DiD fashion, these average outcomes also interacted with *Surname* to account for different levels of investment in innovation across language groups.

All other controls remain identical to the ones previously listed in Section 6.1.1.

6.1.3. Difference in differences: Firm fixed effects

None of the models discussed above can fully account for time-invariant firm heterogeneity. With firm fixed effects, we can then study registration holding firm fixed effects as fixed, under the caveat that these results will be valid for the subset of firms which have changed the language group of their CEO at least once. We then develop the following model:

$$\text{Registered}_{it} = \alpha_i + \text{Surname}_{it}\theta + \text{Surname}_{it} * \text{Post}_t\delta + \text{Post}_t\zeta + \text{Surname}_{it} * \text{Trends}'_{mt}\phi + \text{Trends}'_{mt}\psi + \text{Controls}'_{it}\zeta + \text{Financials}'_{it-1}\eta + e_{it} \quad (6.5)$$

This setting exploits the discontinuity in the introduction of the policy to study whether firms which applied for startup status later in their life also had an Italian-assigned CEO (Eq. (6.5)). Year-fixed effects would be collinear with firm age-fixed effects, so we keep the latter. The initial year of registration is then absorbed by the firm fixed effects, while the *Post* term now denotes the post-policy period. Average variation in outcomes for all other firms in municipality m and year t is controlled for again by our set of controls *Trends*.

6.2. Results

We now discuss the link between language assignment and policy access. Standard errors are always clustered at the level of the firm for every model presented in this section. Table 2 illustrates the language assignment and policy adoption connection starting with the baseline specification (Columns 1 to 4) and the standard saturated model (Columns 5 to 9) from Section 6.1.1.

Columns (1) and (2) offer baseline results for the linguistic assignment of the CEO and the board of directors, omitting all fixed effects except for the year intercepts. The effect is negative and significant for both specifications, featuring nearly the same magnitude. Firms with a German-named CEO are 4% less likely to apply for the policy, and firms with a full board of German-named directors are also 4.2% less likely to apply. Filtering the sample in column (3) so that only firms which are operating in industries which have produced innovative start-ups in the region,²⁸ we are met with a larger -12% effect. Looking, instead, at sole proprietors only in column (4), we find a smaller 2.5% effect.

There is much variance in these baseline results until we shift to the saturated model in the following columns. Column (5) features results from the saturated fixed effects model (Section 6.1.1), holding municipality, industry, year and all their interactions as fixed, and adding all remaining firm-level controls.²⁹ The estimated effect of language assignment remains negative and statistically significant at the 0.001 level but is nearly halved when compared to the baseline model. In fact, firms with CEOs with an etymologically German surname are found to be 2.6% less likely to apply for start-up status, suggesting that time-variant local-industry factors account for around 40% in the variation in policy adoption. Variation in the language composition of the board of directors (Column 6) leads to a similar -2.4% effect.

²⁸ See our discussion in Appendix D for further details.

²⁹ These controls are omitted not only for space reasons but also because, in our tests, their inclusion did not affect the linguistic assignment coefficient, suggesting that they are conditionally orthogonal to it.

²⁷ See Appendix D for more details.

Table 2
Ethnolinguistic assignment and policy adoption, post-policy.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Registered as start-up								
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
CEO surname: German origin	-0.040*** (0.004)		-0.122*** (0.012)	-0.025*** (0.005)	-0.026*** (0.006)		-0.045*** (0.017)	-0.017** (0.008)	
BoD surname %: German origin		-0.042*** (0.004)				-0.024*** (0.006)			
CEO surname ×2013									-0.003 (0.003)
CEO surname ×2014									-0.013** (0.005)
CEO surname ×2015									-0.033*** (0.008)
CEO surname ×2016									-0.020** (0.009)
CEO surname ×2017									-0.035*** (0.009)
CEO surname ×2018									-0.037*** (0.009)
CEO surname ×2019									-0.033*** (0.011)
Observations	24,265	24,289	7533	14,015	21,637	21,637	6362	11,851	21,637
Adjusted R-squared	0.014	0.013	0.053	0.008	0.093	0.093	0.217	0.041	0.093
Fixed effects	Year	Year	Year	Year	Saturated	Saturated	Saturated	Saturated	Saturated
Other controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Filter	None	None	ATECO	S.P.	None	None	ATECO	S.P.	None

Notes: SE clustered by firm in parentheses. Other controls: see Section 6.1.1.
*p<.05; **p<.01; ***p<.001.

Table 3
Ethnolinguistic assignment and policy adoption, Difference in differences.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Registered as start-up							
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
CEO surname × Post	-0.027*** (0.006)		-0.046*** (0.014)	-0.019** (0.008)	-0.020*** (0.004)		-0.052*** (0.011)	-0.020*** (0.006)
CEO surname: German origin	-0.048 (0.042)		-0.239 (0.218)	-0.037 (0.053)	0.033 (0.025)		0.007 (0.087)	0.016 (0.040)
BoD surname % × Post		-0.024*** (0.007)				-0.021*** (0.004)		
BoD surname %: German origin		-0.072 (0.051)				0.019 (0.026)		
Post					0.013*** (0.004)	0.013*** (0.004)	0.030*** (0.011)	0.016*** (0.006)
Firm density (ln)					-0.016*** (0.005)	-0.016*** (0.005)	-0.066*** (0.017)	-0.009 (0.007)
Observations	28,636	28,636	9289	15,807	29,929	29,929	10,144	17,163
Adjusted R-squared	0.105	0.104	0.260	0.056	0.881	0.881	0.878	0.887
Fixed effects	Saturated	Saturated	Saturated	Saturated	Firm-year	Firm-year	Firm-year	Firm-year
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Filter	None	None	ATECO	S.P.	None	None	ATECO	S.P.

Notes: SE clustered by firm in parentheses. Other controls: see Sections 6.1.2 and 6.1.3.
*p<.05; **p<.01; ***p<.001.

Fully disentangling the effects of territorial and industry factors correlated with language is a complicated exercise, but in Appendix G, Table 8, we now introduce all fixed effects in a stepwise fashion, showing that industry and municipality effects are nearly collinear, as the policy effect goes down to -2.4% and -2.6% after controlling for either municipality-year or industry-year cells.

The inclusion of industry controls strongly moderates the effect of filtering the sample by ATECO codes or sole proprietors. In the former case (Column 7), the residual effect of language assignment is only -4.5% for this group of firms, providing an upper bound for the language assignment effect. In the latter case (Column 8), the language effect amounts to -1.7%. In any case, these last results should be taken with a grain of salt because both filters might be endogenous

to language assignment, but the reduced coefficient variance suggests that much of the variation in the coefficients displayed in the baseline models is absorbed by our saturated f.e. strategy.

Finally, we include heterogeneous year-surname coefficients in Column (9), updating the model from Column (5) to account for the possibility that the language assignment effect changes year by year. Our results suggest that, apart from a transitory adjustment period in 2013 and 2014, the effect of language assignment has remained stable at around -3% since 2015.

In Table 3, we move to the DiD models, both with saturated fixed effects (Columns 1 to 4) and firm f.e. (Columns 5 to 8). In both cases, the coefficient of interest is now the one given by the interactions between surname and the post-treatment years *Surname × Post*.

Column (1) begins with the CEO assignment effect in the saturated DiD model from Section 6.1.2. The estimated effect is statistically significant and equals -2.7% , only 0.01 points smaller than the effect estimated previously in Table 2 for the saturated model. Similar magnitude and significance can be noted when looking at BoD composition (-2.4% , Column 2). Columns (3) and (4) filter the sample by ATECO codes and sole proprietors, respectively, finding again an effect that is comparable to the one estimated earlier. These results suggest that our estimates from the standard saturated model were already absorbing much of the existing unobserved differences among language groups which could have been correlated with policy take-up.

Column (5) replaces all fixed effects with two-way firm-years of activity fixed effects (Section 6.1.3), exploiting the discontinuity in policy introduction and variation in language assignment to show results which are valid for all firms which were not born as startups. Our estimates indicate that among this group of firms, those with German-assigned CEOs were 2% less likely to apply for the policy. Column (6) exploits variation in the board's composition, finding a similar effect of -2.1% again. Filtering the sample by ATECO codes (column 7) also leads to results that are not too dissimilar from our previous ones (-5.2%). Looking at sole proprietors in column (8), the inclusion of firm fixed effects makes the language assignment coefficient statistically indistinguishable from the one estimated in column (5). The firm fixed effect allows us to include some municipality-level controls in the mode, such as firm density: in any case, their inclusion does not affect our linguistic assignment estimates.

All these results point to a strong connection between language assignment and policy adoption. The fact that the firm fixed effects model, which gets rid of municipality and industry intercepts, produces the same results further corroborates our findings. The robustness of these results to time-variant and invariant municipality and industry shocks, along with firm fixed effects, suggests that these differences cannot be explained by idiosyncratic municipal and firm factors, provincial policy shocks or simply by proximity to Italy or German-speaking countries. In other words, even in majority-German-speaking municipalities, it is CEOs from the Italian surname group that tend to register their firms as innovative start-ups.

Our three strategies allow us to rule out that differences between the two provinces might have affected our results. Our approach also suggests that our results are valid net of idiosyncratic municipality characteristics, including the proportion of each linguistic group, and general differences in economic performance and industrial structure. Time variance for any of these factors is also accounted for by the saturated fixed effects models. As variation between linguistic groups is also absorbed in the differences in difference model, showing also that pre-trends are in most cases superfluous, we find it difficult to believe that factors other than linguistic assignment could still have played a role in policy adoption.

Still, it could be argued that eligible German-assigned firms might have access to different types of funding. This could be the case if firms within the same municipalities and industries had preferential access to funding from other sources, and if, to firms, funding sources were perfect substitutes. In Appendix G, Table 10, we then perform a robustness check on the saturated model (Section 6.1.1) by introducing controls for firms' access to EU-ERDF funding (column 1) and provincial grants for start-up creation (column 2). Both sets of controls are included in column (3).

If access to these funding sources were to be correlated with language, we would see significant variation in the language assignment coefficient. Instead, the influence on the CEO surname coefficient is marginal at best. The estimated language assignment effect is around -2.4% when both sets of controls are included, which is only 0.2 percentage points larger than the effect previously estimated. The overall effect of grant access to policy adoption is also ambiguous: while access to ERDF grants is negatively correlated with registration, the positive

coefficient of all other funding controls points at complementarities between start-up registration and other sources of firm funding.

7. Discussion and conclusions

The main finding of this paper is that ethnolinguistic diversity has the potential to jeopardise the take-up of a policy, penalising minority groups on the basis of their ethnolinguistic membership. We have shown that the linguistic divide, proxied by the surnames of firms' CEOs, is crucial for determining voluntary participation in a policy, even among firms based in the same country. In particular, young firms with a CEO belonging to a German-speaking group are significantly less likely to access the benefits they are eligible for.

As the language distribution is geographically determined, we develop a strategy that can keep time-variant higher-order factors of institutional, political and industrial nature, which might be endogenous to policy adoption, in check. After controlling for these factors, we find no evidence of differences in economic performance attributable to ethnolinguistic groups during the pre-policy years. The fact that these asymmetries persist year-by-year within the same combinations of municipalities and industry classes, and even after absorbing unobserved differences among language groups in a DiD setting, suggests that residual differences in registration behaviour can only be attributed to exogenous allocation to an ethnolinguistic group and not to differences in economic performance between the two groups. These same asymmetries are also observable in local news outlets, as the German-language press appears to be far less attentive to the policy than the Italian-language one. Similarly, evidence from a survey on residents' policy awareness suggests that native German speakers are more likely to report superficial to no knowledge of national policies than their Italian-speaking counterparts.

We do not believe that the cause of these asymmetries is to be attributed to the failings of regional authorities or language comprehension issues. Our survey data analysis seems to suggest that this is instead a matter of network embeddedness coalescing into a source of information asymmetries, as differences in policy awareness, which re-emerge for a set of flagship national policies, disappear completely once we control for the degree of interconnectedness with individuals from the opposite language group. Our analysis of text data from local news outlets supports this view, suggesting that German-speaking entrepreneurs are less integrated into the Italian-speaking entrepreneurial network and are thus less attentive to news from the 'centre'. In turn, they are also less likely to share relevant information within their network, increasing the costs of access to information and professional expertise embedded in the Italian-speaking network. These costs could be far from negligible for small and fragile firms such as start-ups.

Qualitative interviews conducted with local stakeholders (reported in full in Appendix H) provided additional insights into these dynamics, supporting the idea that this is a matter of information and network embeddedness. In the words of the stakeholders we interviewed, a 'German-speaking CEO is more inclined to seek help from the province rather than from the national arena'. Accordingly, German-speaking entrepreneurs 'leave the money on the table' simply because they have, at best, only superficial knowledge of the policy and might even believe it is not intended for them, especially if they are 'overwhelmed with information and do not know how to process it'.

These findings should not suggest that these information asymmetries may be the sole determinant of policy adoption. Other firm-level factors might still be at play, as our results do not necessarily rule out that other idiosyncratic cultural and institutional factors might also intervene or be themselves sources of information asymmetry. These could include trust in national politics and policies, which may also constitute a reasonable motivation for the lower press consideration. The value of our contribution resides, instead, in showing that these information asymmetries do exist and that, even when linguistic groups are similar enough (or are rendered so via appropriate identification

strategies), cultural proximity with the ‘centre’ remains a source of information and opportunity advantage.

It follows that our results could be externally valid in other multilingual contexts in which it is difficult to rule out that linguistic asymmetries are not linked with other confounding factors. Trentino-Alto Adige is an ideal case study because of its economic and demographic homogeneity, which, net of geography, makes language and attachment to the national network the most meaningful divide between the two groups, all while national policies remain exogenous to the region. As such, our results are relevant to settings in which minority languages correlate with different socio-economic conditions, which can also affect access to information and economic opportunities, such as countries that have a larger share of migrants who could be addressed by targeted promotion of policy programmes and business support schemes. Similarly, our results are relevant to other plurilingual countries – e.g. Belgium and Switzerland – where multiple languages are simultaneously official national languages and in which policies might endogenously take into account the presence of multiple language groups.

Finally, our results carry some relevant policy implications. Importantly, we do not suggest that policymakers should differentiate the provision of their policy across linguistic groups: this would heighten language-based inequalities in the long run (Desmet et al., 2012). On the contrary, the results of this paper encourage the creation of targeted information campaigns that can lead to more inclusive policy adoption. Furthermore, our results have implications for non-multilingual contexts inasmuch as they underline the importance of considering how network embeddedness influences access to information for insulated groups of potential beneficiaries.

CRediT authorship contribution statement

Michele Cantarella: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nicolò Fraccaroli:** Conceptualization, Data curation, Investigation, Writing – original draft, Writing – review & editing. **Roberto Volpe:** Conceptualization, Data curation, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

The authors would like to thank Janina Engel, Davide Viviano, and three anonymous referees for their invaluable comments and suggestions. We also would like to thank Antonio Accetturo for sharing the PAT-PAB dataset with us, and Bodo Steiner for allowing us to use the policy awareness data from Trentino Alto Adige. We are also grateful to participants at the 2021 Rethinking Clusters International Conference, the Helsinki GSE Industrial Organisation Seminar Series, the Labore seminar series, and the 2022 Asian and Australasian Society of Labour Economics Conference for their comments and suggestions. All remaining errors are ours. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for

Reconstruction and Development/World Bank and its affiliated organisations, or those of the Executive Directors of the World Bank or the governments they represent.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.respol.2024.105004>.

References

- Abel, A., Vettori, C., Forer, D., 2012. Learning the neighbour's language: The many challenges in achieving a real multilingual society—the case of second language acquisition in the minority–majority context of south tyrol. *Eur. Yearb. Minor. Issues Online* 9 (1), 271–303.
- Accetturo, A., 2022. Subsidies for innovative start-ups and firm entry. *Ind. Corp. Chang.* 31 (5), 1202–1222.
- Accetturo, A., Barboni, G., Cascarano, M., Garcia-Appendini, E., 2023. The role of culture in firm-bank matching. *J. Financ. Intermediation* 53, 101018.
- Accetturo, A., Cascarano, M., Degasperi, P., Modena, F., 2019. The effects of common culture and language on economic exchanges: evidence from tourist flows. *Reg. Stud.* 53 (11), 1575–1590.
- Acz, Z.J., Audretsch, D.B., 1987. Innovation, market structure, and firm size. *Rev. Econ. Stat.* 69 (4), 567.
- Albanese, G., Bronzini, R., Lavecchia, L., Soggia, G., 2019. Italian Regional Policies for Innovative Start-Ups. *Questioni di Economia e Finanza (Occasional Papers) No. 511*, Banca d'Italia.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. *J. Econ. Growth* 8 (2), 155–194.
- Alesina, A., Giuliano, P., 2015. Culture and institutions. *J. Econ. Lit.* 53 (4), 898–944.
- Alesina, A., La Ferrara, E., 2005. Ethnic diversity and economic performance. *J. Econ. Lit.* 43 (3), 762–800.
- Andersen, K.V., 2013. The problem of embeddedness revisited: Collaboration and market types. *Res. Policy* 42 (1), 139–148.
- Andrews, D., Criscuolo, C., Menon, C., 2014. Do Resources Flow to Patenting Firms? Cross-Country Evidence from Firm Level Data. *OECD Economics Department Working Papers No. 1127*, Organisation for Economic Co-Operation and Development (OECD).
- Angerer, S., Glätzle-Rützler, D., Lergetporer, P., Sutter, M., 2016. Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *Eur. Econ. Rev.* 90, 254–264. Social identity and discrimination.
- Autor, D.H., 2013. The “task approach” to labor markets: an overview. *J. Labour Mark. Res.* 46 (3), 185–199.
- Banca d'Italia, 2021. *L'economia delle Province autonome di Trento e di Bolzano. Economie Regionali - Rapporto annuale N. 4*, Banca d'Italia.
- Baumol, W.J., 1990. Entrepreneurship: Productive, unproductive, and destructive. *J. Polit. Econ.* 98 (5), 893–921.
- Baumol, W.J., 2004. Entrepreneurial enterprises, large established firms and other components of the free-market growth machine. *Small Bus. Econ.* 23 (1), 9–21.
- Bazzi, S., Gaduh, A., Rothenberg, A.D., Wong, M., 2019. Unity in diversity? How intergroup contact can foster nation building. *Amer. Econ. Rev.* 109 (11), 3978–4025.
- Beach, B., Jones, D.B., 2017. Gridlock: Ethnic diversity in government and the provision of public goods. *Am. Econ. J. Econ. Policy* 9 (1), 112–136.
- Breschi, S., Lassébie, J., Menon, C., 2018. A Portrait of Innovative Start-Ups Across Countries. *OECD Science, Technology and Industry Working Papers No. 2018/02*, Organisation for Economic Co-Operation and Development (OECD).
- Buldyrev, S.V., Growiec, J., Pammolli, F., Riccaboni, M., Stanley, H.E., 2007. The growth of business firms: Facts and theory. *J. Eur. Econom. Assoc.* 5 (2–3), 574–584.
- Buldyrev, S.V., Growiec, J., Pammolli, F., Riccaboni, M., Stanley, H.E., 2020. *The Rise and Fall of Business Firms: A Stochastic Framework on Innovation, Creative Destruction and Growth*. Cambridge University Press.
- Calvino, F., Criscuolo, C., Menon, C., 2015. *Cross-Country Evidence on Start-Up Dynamics*. *OECD Science, Technology and Industry Working Papers No. 2015/06*, Organisation for Economic Co-Operation and Development (OECD).
- Cantarella, M., Fraccaroli, N., Volpe, R., 2023. Does fake news affect voting behaviour? *Res. Policy* 52 (1), 104628.
- Chen, M.K., 2013. The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *Amer. Econ. Rev.* 103 (2), 690–731.
- Coia, V., Boschi, I., Trombetta, F., Cavulli, F., Montinaro, F., Destro-Bisol, G., Grimaldi, S., Pedrotti, A., 2012. Evidence of high genetic variation among linguistically diverse populations on a micro-geographic scale: a case study of the Italian Alps. *J. Hum. Genet.* 57 (4), 254–260.
- Cooke, P., Clifton, N., Oleaga, M., 2005. Social capital, firm embeddedness and regional development. *Reg. Stud.* 39 (8), 1065–1077.

- Criscuolo, C., Gal, P.N., Menon, C., 2014. The Dynamics of Employment Growth: New Evidence from 18 Countries. OECD Science, Technology and Industry Policy Papers No. 14, Organisation for Economic Co-Operation and Development (OECD).
- Dahlberg, M., Edmark, K., Lundqvist, H., 2012. Ethnic diversity and preferences for redistribution. *J. Polit. Econ.* 120 (1), 41–76.
- de Chaisemartin, C., D'Haultfœuille, X., 2018. Fuzzy differences-in-differences. *Rev. Econom. Stud.* 85 (2), 999–1028.
- Decker, R.A., Haltiwanger, J., Jarmin, R.S., Miranda, J., 2016. Where has all the skewness gone? The decline in high-growth (young) firms in the U.S. *Eur. Econ. Rev.* 86, 4–23.
- Desmet, K., Ortuño-Ortín, I., Wacziarg, R., 2012. The political economy of linguistic cleavages. *J. Dev. Econ.* 97 (2), 322–338.
- Desmet, K., Ortuño-Ortín, I., Wacziarg, R., 2016. Linguistic cleavages and economic development. In: Ginsburgh, V., Weber, S. (Eds.), *The Palgrave Handbook of Economics and Language*. Springer, pp. 425–446.
- Desmet, K., Wacziarg, R., 2020. The cultural divide. *Econ. J.* 131, 2058–2088.
- Eesley, C., Eberhart, R., Skousen, B., Cheng, J., 2018. Institutions and entrepreneurial activity: The interactive influence of misaligned formal and informal institutions. *Strategy Sci.* 3 (2), 393–407.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. *Q. J. Econ.* 133 (4), 1645–1692.
- Gallo, J.L., Plunket, A., 2020. Regional gatekeepers, inventor networks and inventive performance: Spatial and organizational channels. *Res. Policy* 49 (5), 103981.
- Galor, O., Özak, Ö., Sarid, A., 2020. Linguistic traits and human capital formation. *AEA Pap. Proc.* 110, 309–313.
- Gilsing, V., Nootboom, B., Vanhaverbeke, W., Duysters, G., van den Oord, A., 2008. Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Res. Policy* 37 (10), 1717–1731, Special Section Knowledge Dynamics out of Balance: Knowledge Biased, Skewed and Unmatched.
- Granovetter, M.S., 1973. The strength of weak ties. *Am. J. Sociol.* 78 (6), 1360–1380.
- Guiso, L., Sapienza, P., Zingales, L., 2008. Social capital as good culture. *J. Eur. Econ. Assoc.* 6 (2–3), 295–320.
- Guiso, L., Sapienza, P., Zingales, L., 2015. Corporate culture, societal culture, and institutions. *Amer. Econ. Rev.* 105 (5), 336–339.
- Hall, P.A., Soskice, D., 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford University Press.
- Haltiwanger, J., Jarmin, R.S., Miranda, J., 2013. Who creates jobs? Small versus large versus Young. *Rev. Econ. Stat.* 95 (2), 347–361.
- Istat/MISE, 2018. *Startup Survey 2016: The First Survey on Innovative Startups in Italy*. Istituto nazionale di statistica.
- Klenow, P.J., Li, H., 2020. *Innovative Growth Accounting*. NBER Working Paper 27015, National Bureau of Economic Research (NBER).
- La Porta, R., Lopez-de Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. *J. Law Econ. Organ.* 15 (1), 222–279.
- Larson, J.M., Lewis, J.I., 2017. Ethnic networks. *Am. J. Political Sci.* 61 (2), 350–364.
- Li, Y., Zahra, S.A., 2012. Formal institutions, culture, and venture capital activity: A cross-country analysis. *J. Bus. Ventur.* 27 (1), 95–111.
- Menon, C., DeStefano, T., Manaresi, F., Soggia, G., Santoleri, P., 2018. The evaluation of the Italian “Start-up Act”. OECD Science, Technology and Industry Policy Papers 54, OECD Publishing, Organisation for Economic Co-Operation and Development (OECD).
- MISE, 2019. *The Italian Startup Act. Italy’s Policy Framework to Support Innovative Startups: Executive Summary*. Policy Summary 07/2019, Italian Ministry of Economic Development.
- MISE, 2021. *Relazione Annuale al Parlamento sullo stato di attuazione e l’impatto delle policy a sostegno di startup e PMI innovative*. report, Italian Ministry of Economic Development.
- Montalvo, J.G., Reynal-Querol, M., 2021. Ethnic diversity and growth: Revisiting the evidence. *Rev. Econ. Stat.* 103 (3), 521–532.
- Murray, F., Aghion, P., Dewatripont, M., Kolev, J., Stern, S., 2016. Of mice and academics: Examining the effect of openness on innovation. *Am. Econ. J. Econ. Policy* 8 (1), 212–252.
- Nifo, A., Vecchione, G., 2014. Do institutions play a role in skilled migration? The case of Italy. *Reg. Stud.* 48 (10), 1628–1649.
- North, D.C., 1990. *Institutions, Institutional Change and Economic Performance*. Political Economy of Institutions and Decisions, Cambridge University Press.
- OECD, 2020a. *An Insight into the Innovative Start-Up Landscape of South Tyrol*. OECD Local Economic and Employment Development (LEED) Papers No. 2020/07, Organisation for Economic Co-Operation and Development (OECD).
- OECD, 2020b. *An Insight into the Innovative Start-Up Landscape of Trentino*. OECD Local Economic and Employment Development (LEED) Papers No. 2020/09, Organisation for Economic Co-Operation and Development (OECD).
- OECD, 2020c. *Start-Ups in the Time of COVID-19: Facing the Challenges, Seizing the Opportunities*. OECD Policy Responses to Coronavirus (Covid-19) 13 May 2020, Organisation for Economic Co-Operation and Development (OECD).
- Schot, J., Steinmueller, W.E., 2018. Three frames for innovation policy: R&D, systems of innovation and transformative change. *Res. Policy* 47 (9), 1554–1567.
- Spolaore, E., Wacziarg, R., 2016. Ancestry, language and culture. In: Ginsburgh, V., Weber, S. (Eds.), *The Palgrave Handbook of Economics and Language*. Palgrave Macmillan, pp. 174–211.
- Sutter, M., Angerer, S., Glätzle-Rützler, D., Lergetporer, P., 2018. Language group differences in time preferences: Evidence from primary school children in a bilingual city. *Eur. Econ. Rev.* 106, 21–34.
- Vettori, C., 2016. The long way to bilingualism: the peculiar case of multilingual South Tyrol. *Int. J. 21st Century Educ.* 3 (2), 25.
- Vettori, C., Abel, A. (Eds.), 2017. *KOLIPSI II. Gli studenti altoatesini e la seconda lingua: indagine linguistica e psicosociale. / Die Südtiroler SchülerInnen und die Zweitsprache: eine linguistische und sozialpsychologische Untersuchung*. Eurac Research, Bolzano/Bozen.
- Vettori, C., Abel, A., Wisniewski, K. (Eds.), 2012. *KOLIPSI. Gli studenti altoatesini e la seconda lingua: indagine linguistica e psicosociale. / Die Südtiroler SchülerInnen und die Zweitsprache: eine linguistische und sozialpsychologische Untersuchung*. Eurac Research, Bolzano/Bozen.
- Vettori, C., Colobo, S., Abel, A., 2021. Multilingualism in South Tyrol: between old fears and new challenges. *Z. Interkulturellen Fremdsprachenunterricht* 1 (26), 223–249.
- Wang, C., Steiner, B., 2015. Can ethno-linguistic diversity explain cross-country differences in social capital?: A global perspective. *Econ. Rec.* 91 (294), 338–366.
- Webb, J.W., Khoury, T.A., Hitt, M.A., 2020. The influence of formal and informal institutional voids on entrepreneurship. *Entrepreneurship Theory Pract.* 44 (3), 504–526.
- Williams, N., Vorley, T., 2015. Institutional asymmetry: How formal and informal institutions affect entrepreneurship in Bulgaria. *Int. Small Bus. J.* 33 (8), 840–861.
- Wooldridge, J., 2021. Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. Mimeo.