Politecnico di Torino

Master's Degree in Engineering and Management

Tesi di Laurea Magistrale

Female Entrepreneurship and Support Systems in the North-West of Italy

Statistical Analysis on the Impact of Caregiving Services and Female Foreign Labour on the Foundation of Female-led Innovative Startups

Relatrice: Prof.ssa Anna D'Ambrosio

Candidata: Melissa Mani

Anno accademico 2023/2024

Abstract

This thesis examines the impact of female immigrants on the foundation of new female-led innovative startups, particularly in relation to local care services, by developing a linear regression model. Starting from the work of Patricia Cortés, the study argues that low-skilled female immigrants significantly contribute to reducing domestic and caregiving responsibilities, therefore enabling high-skilled native women to pursue entrepreneurial careers and potentially increasing the rate of female-led innovative startups.

The research is grounded in a comprehensive literature review that exploits the theoretical frameworks regarding the dynamics of innovative startups and female entrepreneurship. At first, it highlights the determinants enabling the founding of new innovative businesses on a general level, and then within the context of the North-West of Italy.

Afterwards, a linear regression model with multiple regressors is constructed to analyse provincial-level data on female-led startups from the databases of AIDA and the Register of Innovative Startups, foreign population statistics, and the availability of local childcare services from the National Institute of Statistics. The model controls for demographic and economic variables that might also impact the foundation rate of startups.

The findings provide significant insights for developing strategies to promote female-led innovation in startups and support a diverse entrepreneurial ecosystem.

Contents

Index of Figures and Tables

1. Introduction

The role of female entrepreneurship in driving economic innovation and societal progress is increasingly recognised worldwide. Nevertheless, despite notable advancements, women continue to be underrepresented in the labour market, particularly in leadership positions, within the STEM fields, and in the entrepreneurial ecosystem.

Italy, a country traditionally characterised by a low female labour participation rate and established traditional gender roles, presents unique challenges for female entrepreneurs. The North-West of Italy, which includes innovative entrepreneurial ecosystems such as Milano and Torino, presents a dynamic and robust environment for innovative startups. However, women in this region face barriers such as limited access to fundings, lack of supportive and targeted policies, and the dual responsibility of caregiving and professional work. These constraints, together with the relatively high costs and limited availability of public care services, often discourage and obstacle women from pursuing entrepreneurial careers.

Considering these challenges, this thesis explores the role of immigrant women and of caregiving services in increasing the presence of female-led innovative startups in the North-West of Italy. Building on the research of Patricia Cortés, this study argues that low-skilled female immigrants, by taking on domestic and caregiving responsibilities, enable high-skilled native women to devote more time to entrepreneurial activities. Low-skilled labour provided by foreign women, thus, acts as a private substitute for the public welfare services, creating advantageous conditions for the founding of female-led innovative startups.

The study provides a comprehensive literature review in Chapter 2, covering the theoretical background on female entrepreneurship and innovative startups, as well as the specific factors influencing the entrepreneurial landscape in Europe and North-West Italy. It delves into the roles of personal attitudes, regional dynamics, and entrepreneurial ecosystems in driving female entrepreneurship in innovative contexts, with particular focus on the interaction between low-skilled female immigrants and native women entrepreneurs.

The third chapter introduces the empirical methodology used in this study. A linear regression model is developed to analyse provincial-level data, assessing the impact of caregiving services and female immigrant labour on the formation of innovative startups led by women.

Afterwards, in Chapter 4 it is explained the construction of the dataset with observations across provinces starting from 2013 to 2022, including economic, demographic, services and immigrants' data obtained by ISTAT – the National Institute of Statistics – and startup data deriving from a merge between the official Register of the Innovative Startups and the AIDA database. In this chapter, the hypothesis of the three analysed linear regression models are described, followed by a descriptive analysis outlining how the variables are integrated and distributed to ensure robustness in the statistical analysis.

In the end, Chapter 5 presents the results of the empirical models, highlighting key findings and their implications for female entrepreneurship and, especially, for the foundation rate of innovative startups led by women. It is, finally, followed by Chapter 6 that discusses the limitations of the study and potential directions for future research.

2. Literature Review

2.1 Innovative Startups and Female Entrepreneurship

This chapter delves into the arising field of female-led innovative startups in the North-West of Italy, exploring the complex interplay between gender, immigration, and entrepreneurial dynamics.

The review begins with an overview of innovative startups and female entrepreneurship across Europe, setting the context to then examine the key factors influencing the creation and success of female-led startups, with a particular focus on the impact of low-skilled female immigrants on these businesses, highlighting how this demographic contributes to the startup ecosystem. Following this, the review offers a detailed overview of the Italian landscape for innovative startups, with a specific emphasis on the role of female entrepreneurs within this scenario. This comprehensive examination aims to provide an expanded understanding of how these interconnected elements shape the innovative startup environment in the region.

2.1.1 The Concept of Innovative Startup

The review of the existing literature suggests that there is no unique definition for "Innovative startup". Indeed, a variety of definitions have been assigned by authoritative sources - national and international organisations - and researchers, subject to the scope and age of activity of these entities, their growth rate in terms of investments and R&D activities, and the significance of the innovativeness introduced by technology.

Nevertheless, these definitions present some common elements, outlined as follows, which reflect a consensus among international and national organisations on the critical role of innovative startups in economic development and technological progress.

- <u>Centrality of innovation</u> → introducing a new or significantly improved product, service, or process.

- <u>Technology</u> → making use of advanced technology or technical knowledge.

- <u>Growth potential</u> → high potential for rapid growth and scalability.

- R&D investments → high investments in R&D activities.

- Market disruption → ability to disrupt existing markets and industries.

National definitions often align with broader international frameworks but may include specific criteria relevant to local policies and economic objectives.

With reference to Italy, the national Government has been engaged in the formulation of an all-embracing and consistent legislation since 2012, with the purpose of creating favourable conditions for the establishment and the development of new innovative enterprises with high-technological value and congruent with the needs of all the players involved in the Italian innovation ecosystem. It resulted in the introduction of the definition of innovative startups into the Italian legal system, through the Decree-Law 179/2012, also known as the Italian Startup Act, converted into Law 221/2012.

The legal notion of innovative startup that is used for the scope of this thesis, is set out in Art. 25 of Decree-Law no. 179/2012, paragraph 2, and reported hereby: *"Any companies with shared capital (i.e. limited companies, "società di capitali"), including cooperatives, whose capital shares – or equivalent – are neither listed on a regulated market nor on a multilateral negotiation system."* (Italian Ministry of Economic Development, 2012)

These enterprises must also comply with the following requirements:

1) They are newly established or have been incorporated for less than 5 years (in any case, not before 18 December 2012).

2) They have their headquarters in Italy, or in another EU/EEA Member State provided that they have a production facility or a branch in Italy.

3) They have an annual turnover lower than €5 million.

4) They do not distribute their profits and have not done so in the past.

5) Their mission statement ("oggetto sociale") concerns, predominantly or exclusively, the development, production and commercialisation of innovative products or services with a clear technological component.

6) They are not the result of a company merger or split-up, or of a business or branch transfer.

7) Finally, they meet at least one of the three following innovation-related indicators:

   a) Research and development expenditure corresponds to at least 15% of the higher value between turnover and annual costs (as per the last statement of accounts).

   b) The total workforce includes at least 1/3 of PhDs, PhD students or researchers, or at least 2/3 of the team hold a master's degree.

   c) The company is the owner or licensee of a registered patent (or it has filed an application for an industrial property right) or it owns an original registered software.

It is widely believed that innovative startups have a significant role in driving economic development, technological progress and societal impact. These startups not only generate high-wage employment but also exert competitive pressure on existing enterprises, compelling them to adapt or exit the market. This dynamic, described by Schumpeter as "creative destruction”, enhances productivity by optimising resource allocation within the economy. Empirical research suggests, indeed, a positive relationship between an economy's productivity growth and the rate at which businesses enter and exit the market. The critical role of technological innovation in enhancing a country's growth and productivity underscores the intrinsic link between startups and innovation, positioning them as essential elements in fostering global progress. Schumpeter's work, "The Theory of Economic Development" (1911), established a foundational connection between innovation and economic prosperity, emphasising the key role of innovation in society. Innovative startups encounter an additional liability of novelty as they introduce new products, services, or processes, further challenging their survival, especially when they possess a high-risk profile. This duality of innovative startups as both a unique form of entrepreneurship and a mechanism for innovation development presents a complex challenge for policymakers. On one hand, these firms align with entrepreneurship policies aimed at fostering new ventures that can survive and contribute to job creation and economic growth. On the other, they are influenced by innovation policies designed to enhance industrial competitiveness and societal renewal. The tension arises from the fact that the most innovative startups, which present the greatest potential for growth and impact, also tend to experience the highest failure rates. To guarantee the long-term growth

of the national economy, measures must be taken to enhance the startup ecosystem's knowledge support. This will enable immediate action in response to adverse changes in the startups' dynamics, thereby resulting in more chances for sustainable development.
Additionally, with an increase in the implementation of startups there is a positive dynamic in the implementation of the UN SDGs, with a stable positive relationship characteristic not only for economic, but also for social, environmental, and institutional goals, which significantly increases the importance of startups for the sustainable development of territories and thus, making them increasingly attractive for policy makers.

2.1.2 Overview on European Framework

During the past few years, the European startup ecosystem has experienced considerable growth and dynamism, characterised by many high-growth sectors and several leading innovation hubs. Between 2018 and 2022, Europe saw a substantial rise in early-stage funding activities, particularly in sectors such as Fintech, AI, Big Data, and Healthtech.
Although startups raised less funding globally in 2022 than the previous few years, Europe's funding landscape remained resilient, supported by unicorns and high value exits in cities like London, Berlin, and Amsterdam. The continent's share of global startup funding increased from 18.3% in 2021 to 19.7% in 2022, reflecting, indeed, the modest increase in investment in European startups.
The success of the European startup ecosystem is also strengthened by supportive policies and frameworks at both the national and EU levels. European Union’s organisations initiatives and various national programs provide funding, mentorship, and resources to startups. Furthermore, the regulatory environment is increasingly focused on promoting innovation, with plans to streamline regulations and reduce barriers for startups operating across multiple countries.
Geographically, London continues to lead as Europe’s top startup hub. Nevertheless, the European startup ecosystem is highly interconnected, facilitating the movement of talent and capital across borders. The Startup Heatmap Europe Report 2023 highlights that 31% of founders in Europe are from abroad, and a significant majority of startups maintain a presence in more than one country. This hyper-mobility is crucial for sustaining innovation and collaboration across the continent. Additionally, Europe benefits from a strong network

of accelerators, incubators, and co-working spaces that provide vital support to early-stage startups, according to DEEP Ecosystems.

As follows, a graph showing data from Startup Heatmap Europe 2023 with information about where most founders were born is presented. The graph ranks the cities based on multiple criteria such as the presence of STEM students, unicorn offices, research funding, university founders, accelerator participants, and coworking spaces. Cities such as Milan, Brussels, and Tampere are prominently displayed, demonstrating their contributions to the entrepreneurial world by including robust support systems and active startup communities.

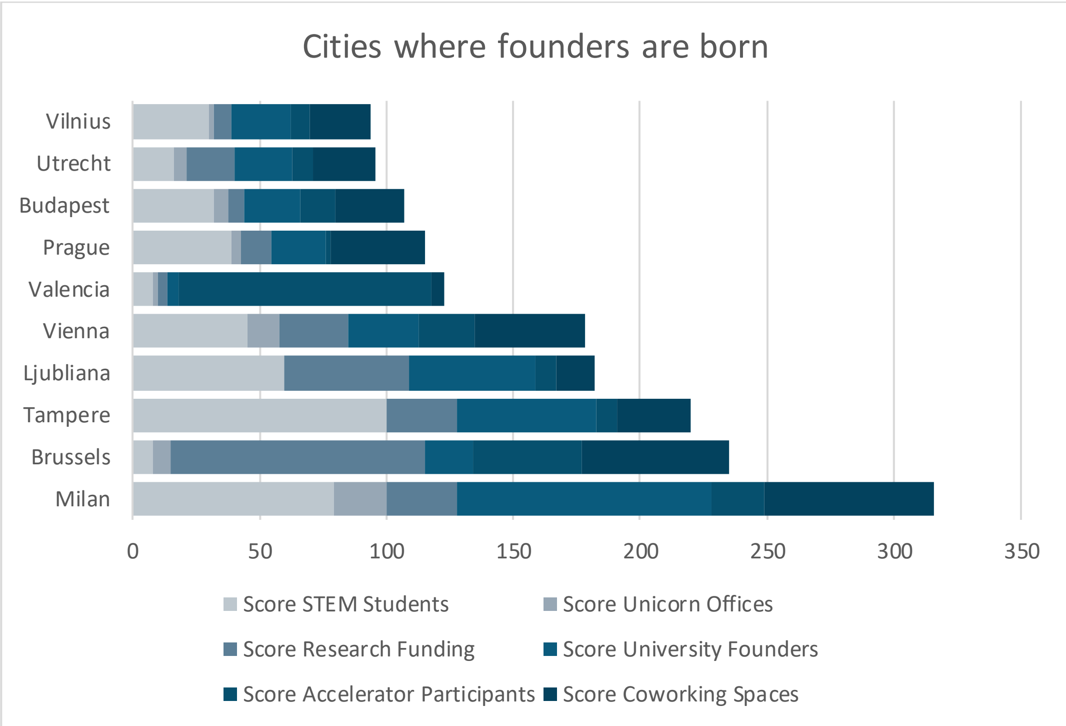

Figure 1: Cities where founders are born (data from Startup Heatmap Europe).

In summary, Europe is a key player in the global startup scene, bringing in talent and capital from around the world and facilitating an environment of innovation and growth through a robust investment landscape, diversity of high-growth sectors, strategic geographic hubs and supportive regulatory system.

2.1.3 Female Entrepreneurship

The Global Entrepreneurship Monitor defines entrepreneurship as the act of starting or running a new business, which brings new products and technologies to the market and turns the ideas of entrepreneurs into tangible goods and services that consumers or other businesses want to buy.

As per InfoCamere - Women's Entrepreneurship Observatory of Unioncamere, this thesis defines female-led enterprises those where the majority of ownership and control is held by women. The degree of female participation is determined by the legal nature of the company, the share of capital held by each member and the percentage of women among the directors or owners or members of the company. Typically, firms with a total participation rate of more than 50% of women are defined as women-led enterprises.

Entrepreneurs, in general, accelerate structural changes in the economy by attracting resources to new sectors and industries while removing resources from sectors that offer goods and services that people no longer need. These kinds of structural adjustments improve productivity, which eventually boosts living standards throughout the economy. Entrepreneurs can also promote social changes, by finding new solutions to society's most pressing challenges. Thus, entrepreneurship plays a crucial role in the process of socio-economic growth.

Accordingly, a group of environmental determinants have been identified as significant in determining the emergence of innovative startups within a region, implying that an ecosystem of interconnected components inhibits entrepreneurship. In 1993, Moore introduced the word "ecosystem" for the first time in management literature to describe the complexity of the external environment. It has been formally defined as "a set of independent actors and factors coordinated in such a way that they enable productive entrepreneurship". This kind of ecosystem revolves around the entrepreneur, and entrepreneurship is its main product.

The definition of the entrepreneurial ecosystem has been described as territory-specific because it includes a range of cultural, social, political, and economic characteristics of a region that may have an impact on the emergence and expansion of novel businesses and may cause local actors to become more dependent on one another in order to produce new value.

It is crucial to look at the statistics regarding female entrepreneurship in Europe to further comprehend the representation of women in the entrepreneurial ecosystem, highlighting trends, progress, and ongoing gaps in various sectors across the continent, before delving deeper into the Italian - specifically North-West Italian - framework at the end of the chapter.

The European business and entrepreneurship ecosystem is far from being equal, diverse and inclusive. According to the EU Startup Monitor 2018 Report supported by the European Commission, a common profile emerges when analysing founders of European startups. The average founder is:

- male (82.8%)
- has a university degree (84.8%)
- is currently 38 years old
- was 35 when founded the business

Though women constitute over 50% of the population in Europe, much fewer are taking the risk of founding their own businesses. The Startup Heatmap Report 2023 reports that the share of female founders or co-founders among the total was 13.2% in 2022, highlighting a drop of -3.7% in comparison to 2021. Even in top ecosystems like Berlin and London, less than one in five business founders - only 18% - are female. While, according to recent statistics of the European Commission, women constitute the majority of one person enterprises, 78%, in Europe and are also over-represented among the smallest businesses in highly competitive, low-margin markets and industries, such as health and social work activities, services and education.

Female entrepreneurship is a source of economic growth that is under exploited, as evidenced by this scenario. Generally, women entrepreneurs are more educated, with higher levels of graduate education - more precisely with a ratio of 1.08 women with respect to men - and fewer women having only secondary or less education compared to men. However, women appear to be more likely than men to start a business due to job scarcity and to make a difference in the world. Job scarcity is, indeed, the predominant reason for starting a business for both genders, with 72.9% of women and 67.2% of men citing it as their primary motivation.

Additionally, there is a lack of VC funding to support female-led startups. In Europe, startups founded solely by women raised just 1.8% of the total capital invested in venture-backed startups in 2023, although startups founded only by women have grown their share of VC deals from 2.7% to 5% between 2008 and 2024, with the United Kingdom, followed by France, being the top European nation for women-founded companies in both deal value and volume.

Despite the increasing optimism surrounding future investments in female-led startups, a significant disparity persists in the level of investments when compared to male-led startups. Female entrepreneurs tend to join accelerator programs 1 to 2 years later than male entrepreneurs and raise less funding when starting under their same conditions. Female-led startups are 5 years behind their male-led businesses, which means that their business models are less likely to grow fast in terms of economic prosperity and impact. Women often partner with men to secure more funding, as a study by Boston Consulting Group (BCG) and SISTA shows that such collaborations can nearly triple the capital raised compared to women-only teams. In a review of five years of investment and revenue data by the same research, it is outlined that startups founded and co-founded by women are more profitable financial businesses in terms of how well they convert a dollar of investment into a dollar of revenue. These startups generated 78 cents for every dollar of capital, compared to less than half that amount—just 31 cents—for startups developed by men.

However, while partnering with men helps women scale their startups and face fewer inequalities, it also results in their male counterparts raising about 30% less than they would with other men. Despite targeted funding initiatives aimed at closing these gaps, mixed-gender teams still face disadvantages in fundraising, with male-dominated teams capturing the majority of opportunities. Although there are occasional successes for female-only teams, overall, gender equality remains a challenge in the European startup ecosystem.

2.2 Determinants of Innovative Startups' Foundation and of Female Entrepreneurship

The establishment of new enterprises serves as a critical measure of business dynamism, highlighting an essential aspect of entrepreneurship within a country: the ability to initiate entirely new businesses. New enterprises are regarded as key drivers of economic growth, owing to their significant contributions to overall job creation and their role in enhancing productivity, which is linked to the dynamic process of firm entry and exit.

Enterprise birth is considered an independent event, impacting only one entity within the broader population of active enterprises. Following OECD's publication, Entrepreneurship at a Glance (2017), this process entails the formation of a new combination of production factors and often results in the registration of a new business entity, depending on the scope and requirements of the business register.

The statistical analysis performed in this thesis aims to assess the variation in the founding rate of female-led innovative startups (dependent variable) in relation to changes in the number of foreign female residents and the availability of caregiving service structures (independent variables) within a specific region. The dependent variable is defined in accordance with the Eurostat-OECD Manual on Business Demography Statistics, which characterises the birth rate for a given reference period as "the number of births as a percentage of the population of active enterprises."

Considering that the aforehand mentioned factors are not the only ones impacting the decision to start an entrepreneurial activity, it is necessary that the model accounts for the other determinants that might also affect the startup foundation rate.

Empirical research investigates two primary aspects of entrepreneurship: the individual traits that facilitate new venture initiation and successful post-entry performance, and the regional dynamics affecting firm entry. These studies have examined how new firm formation influences local economies and how regional characteristics impact these processes (Colombelli A., 2016). Therefore, researchers indicate that entrepreneurs are influenced by their local social, cultural, and economic environments, which either support or constrain their efforts.

As per the literature review reported beforehand, significant disparities persist between male and female entrepreneurs, especially in certain areas such as access to capital, business performance metrics, and representation across various industries. By analysing the reasons behind these gaps, it is possible to further understand the unique paths taken by female entrepreneurs when founding a startup and the barriers they must overcome. Furthermore, it is significant for policy makers to know the elements that impact women's decisions to launch a startup given that encouraging the creation of women-led businesses will not only bring bright business ideas but will also empower the role of women in society and create new sources of wealth.

With the objective of researching female entrepreneurial behaviour in Europe, the Women's Entrepreneurship Report 2022/2023, released by GEM, outlines the statistical disparities between female and male entrepreneurs with regard to the intention to launch a startup: one in six women worldwide report aimed at starting a business in the near future, compared to one in five men; subsequently, one in ten women worldwide are in the early stages of starting a business, compared to one in eight men.

For the purpose of this thesis, it is necessary to extensively examine the enabling factors of entrepreneurship and of the start-up of innovative businesses in general, considering that some of these determinants are gender specific and reflect the societal roles of women.

2.2.1 Personal Attitude toward Entrepreneurship

Among the reasons for choosing an entrepreneurial career, personal interests and objectives have a significant role, especially when dealing with the tech-sector. Women have frequently stated that having a great business concept, wanting to acquire freedom of action, or wanting to improve the world are what drive them to start their own businesses. However, in cases where the decision to start a business is a personal one, it is influenced by a variety of elements. These include the capacity to recognize business potential, self-esteem, personal intentions, proximity to other entrepreneurs and mentors, availability of resources including social support, and a risk-taking mindset.

Women's willingness to take risks is essential to the development of innovative startups and, more significantly, to their very foundation. The literature indicates that although women are generally more risk-averse than males, this cautious attitude can have a major effect on the decision to establish a business. According to researchers Andrea Rey-Martía, Ana Tur Porcar, and Alicia Mas-Tura (2015), these results also suggest a connection between women's attitudes toward business growth and the small size of women-led startups. They conclude that men are more likely to set up new companies given that they are more willing to take risks and pursue business opportunities more aggressively. Conversely, businesswomen adopt a longer-term growth strategy because they are less inclined to take on the risk of expanding a company rapidly. This reasoning arises from a desire to keep a work-life balance, manage and regulate operations in accordance with their expertise and skills, and access resources acceptable for a smaller organisation.

Thus, women's risk aversion—which is frequently fuelled by a lack of confidence in one's own skills and a desire for a work-life balance—can prevent them from deciding to launch their own businesses.

Nonetheless, because they acquire a strong foundation of knowledge prior to embarking on this journey, despite their lack of confidence, women who overcome their risk aversion are more likely to create innovative startups that succeed. Therefore, encouraging female entrepreneurs to embrace risks is essential for both ensuring that more women launch firms and that these ventures are robust and able to deal with all aspects of the entrepreneurial career.

2.2.2 Geographic and Regional Factors

The literature on entrepreneurship has shown that, besides individual traits and personal factors, contextual elements also have an influence on the decision to create a startup.

The local entrepreneurial ecosystem, including organisations such as incubators, accelerators, and research centres, is fundamental to the development and establishment of innovative enterprises, particularly the ones headed by women. These organisations serve as hubs for the transfer of knowledge, giving entrepreneurs access to advanced research, innovative technologies, and opportunities for mentorship - all of which are critical in the early phases of business development. On the other hand, incubators and accelerators provide more specialised support by offering a structured environment for entrepreneurs to grow their business. These organisations usually provide services like office space, business coaching, and seed investment, which are particularly helpful for female entrepreneurs who might face other difficulties including restricted access to networks and financial capital.

An effective regional innovation ecosystem is generally capable of generating knowledge spillovers, fostering academic spinoffs, and cultivating highly specialised human and social capital, including key agents like business angels and venture capitalists. (Colombelli A., et al, 2016).

The knowledge spillover theory of entrepreneurship was conceptualised in 1995 and further developed by economists D. Audretsch, and E. Lehmann in 2005. The theory states that "the context in which decision-making is derived can influence one's determination to become an entrepreneur."

The principle of the spillover concept is that knowledge is not entirely appropriated by those who generated it. Instead, entrepreneurs can profit from this knowledge by choosing to locate their businesses in areas where there are organisations (private, non-profit, government, university, or research institution) that have used their own labour and resources to develop new knowledge that has the potential to be commercialised but decided not to do so for different reasons. This theory highlights the potential for startups - particularly innovative startups - to establish themselves in regions where they can profit from underexploited concepts and technologies that represent a main source of entrepreneurial opportunities (Audretsch et al., 2006; Acs et al., 2013).

It is widely argued within the academics that local connections significantly influence firms' innovation capabilities. Engaging and collaborating with various stakeholders facilitates learning processes and the transfer of expertise, thereby enabling geographical regions to experience elevated rates of new business formation. Concurrently, the Open Innovation - OI - model indicates a decline in the benefits derived from internal R&D expenditures. In contrast to previous decades, innovative firms allocate less funding to internal R&D activities but continue to achieve successful innovation by leveraging knowledge from a diverse range of external sources (D’Ambrosio A. et al., 2016). This reliance on external expertise, particularly through local collaborations, plays a crucial role in fostering innovation and enhancing the rates of new venture creation within these regions.

2.2.3 Institutional and Policy Environment

The significance of new venture creation in driving innovation, employment, and economic growth has been widely recognized, leading to the prominent role of innovative start-ups in both policy discussions and academic discourse in recent years (Autio et al., 2014). Policymakers primarily aim to foster the establishment of startups by alleviating the challenges entrepreneurs encounter when launching new enterprises. However, the outcomes of these policies have not consistently yielded favourable results. Government initiatives designed to encourage entrepreneurship—such as tax incentives, simplified business registration processes, protection of intellectual property rights, and labour market flexibility—have a substantial impact on the rate of new business formation. The regulatory environment can either facilitate or hinder entrepreneurial activity, contingent upon its adaptability and supportiveness towards nascent and small enterprises. Empirical evidence suggests that promoting startup creation without rigorous evaluation of their quality may lead to ineffective public policy (Colombelli et al., 2016). Additionally, the approach of policymakers is transitioning towards the adoption of a comprehensive and interconnected "entrepreneurship policy" framework, by integrating various institutional and infrastructural elements that significantly influence the effectiveness of policy interventions. By leveraging both a national system of entrepreneurship and entrepreneurial ecosystems, such policies can more effectively stimulate the innovative activities of new enterprises (Audretsch D., Colombelli A., Grilli L., Minola T., Rasmussen E., 2020).

2.2.4 Innovation Determinant

Researchers have been involved in a debate regarding the effects of a startup's innovation on its formation, growth and longevity. The literature presents differing perspectives on the ways in which innovation affects the overall performance of a firm.

Theoretical approaches propose both positive and negative effects. Positively, innovation enables entry into the market, increases a firm's market power, lowers costs, and improves capabilities—all of which may promote long-term survival. On the other hand, innovation may also come with more risks, more complicated startup procedures that could be disproportionately impeded by their novelty, and challenges obtaining outside funding, all of which could result in a higher failure rate.

Most of the empirical research indicates a positive correlation between innovation and survival; however, more recent data raises the possibility that this relationship may vary depending on the context, especially for smaller businesses. The belief that innovative ventures are more stable by default is challenged by these data, which suggest that innovation should not be taken as a guarantee of survival during the startup phase.

In addition to having an effect on innovative startups overall, the innovativeness parameter also appears to have an influence on innovative firms led by women specifically. Decisions regarding innovation are, indeed, made based on a variety of internal and external (institutional) considerations that, as previously noted, influence the degree of risk that an innovation could induce to the decision to establish an enterprise and the probability of success.

Using cross-country data from 2008 to 2015 of businesses in 75 economies, D. Audretsch and M. Belitski conducted a study in 2020. By examining the internal dynamics of women-led businesses and comparing how men and women perceive risk differently, the findings indicate that there is no direct connection between risk perception and the propensity towards innovation. Rather, it argues that opportunities for innovation in women-led businesses are mostly dependent on the availability of resources, including internal financial resources and financial flexibility. According to research, innovation is mostly driven by access to resources and capital, and this relationship gets stronger as financial resources rise.

Nevertheless, the analysis on how investors' external perceptions of risk affect women-led businesses' access to capital and the business environment suggests that investors frequently view women-led businesses as riskier investments than those led by men, which has a negative effect on their possibilities to attract capital. In turn, this hinders their potential for innovation and determines the way women are viewed in the business world.

2.2.5 The Status of Women

Regarding innovative and high-tech startups in particular, prior research has indicated that female, young, and foreign-born, entrepreneurs adopt distinct startup processes and, as a result, may have varying effects on a region's startup birth rate. Despite their tendency to run smaller businesses, female entrepreneurs are of particular importance because they are within the categories with the fastest rates of growth.

Contextual, socio-demographic, and individual perceptual characteristics can be used to classify the primary components analysed in most of the existing studies on gender and entrepreneurship. Nonetheless, to adequately contextualise women's qualifications and limits while evaluating the foundation rate of women-led companies, it is important to take into account the following gender-specific factors.

Recently, the role of women in society and the lack of equitable social support from the government and local authorities have emerged as a dominant influential variables in the innovative female-led startups' foundation rate in a geographic area.

Historically women have always been left behind in terms of economic opportunities, to the point where the Economic Participation and Opportunity gender gap is globally closed at only 60.5%, according to the Global Gender Gap Report 2024 from the World Economic Forum.

Besides personal motivation, national and international regulations and the entrepreneurial ecosystem, there are specific aspects to target when analysing women's opportunities and willingness to establish, and then, sustain a business.

Gender-Aware 5 M Model

These aspects are examined and explained by the Gender-Aware 5M Model introduced by Candida Brush, Anne de Bruin, and Friederike Welter in their 2009 paper titled "A gender-aware framework for women's entrepreneurship." It expands on the traditional 3M model including Macro/Meso Environment and Motherhood in addition to Money, Management and Market. Each of these dimensions addresses specific challenges and opportunities that women may face in the entrepreneurial landscape, taking into account gender-specific factors.

MONEY - Access to Finance

Among the gender limitations that women founders encounter, as it was previously reported, getting investments and access to financial resources is one of the most critical ones, especially in slowly changing male-dominant industries. Financial capital is essential for the foundation of startups. Access to venture capital, angel investors, or public funding opportunities directly influences the rate at which new firms can be founded and scaled. Innovative startups, especially, rely on external funding to cover R&D costs and to scale rapidly.

The graph provided by Dealroom.com, a global startup data platform, enables to observe how, in Europe between 2019 and 2023, the proportion of investments and capital raised by startups with at least one female founder differs from that of startups with all male founders.

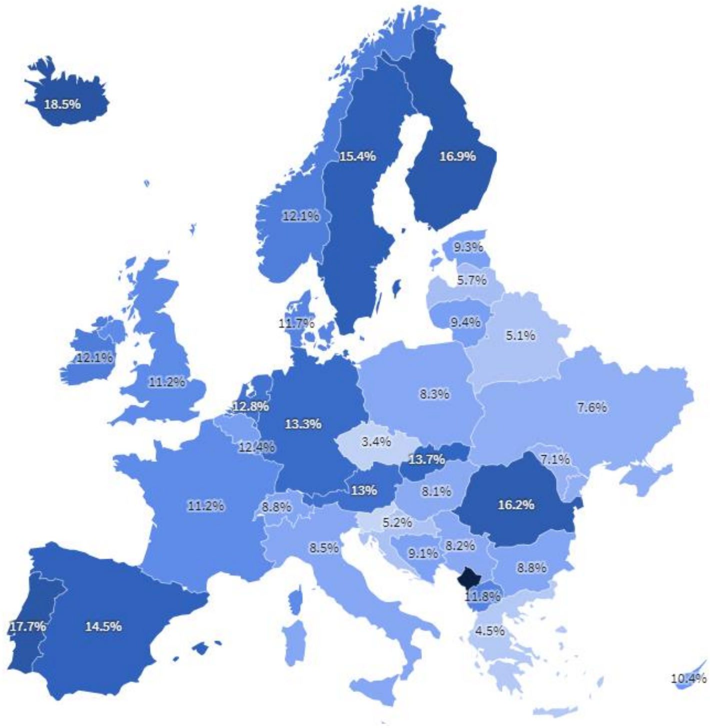

Figure 2: Percentage of rounds raised by female founded startups as % of country total (2019 - 2023) - data from Dealroom.com.

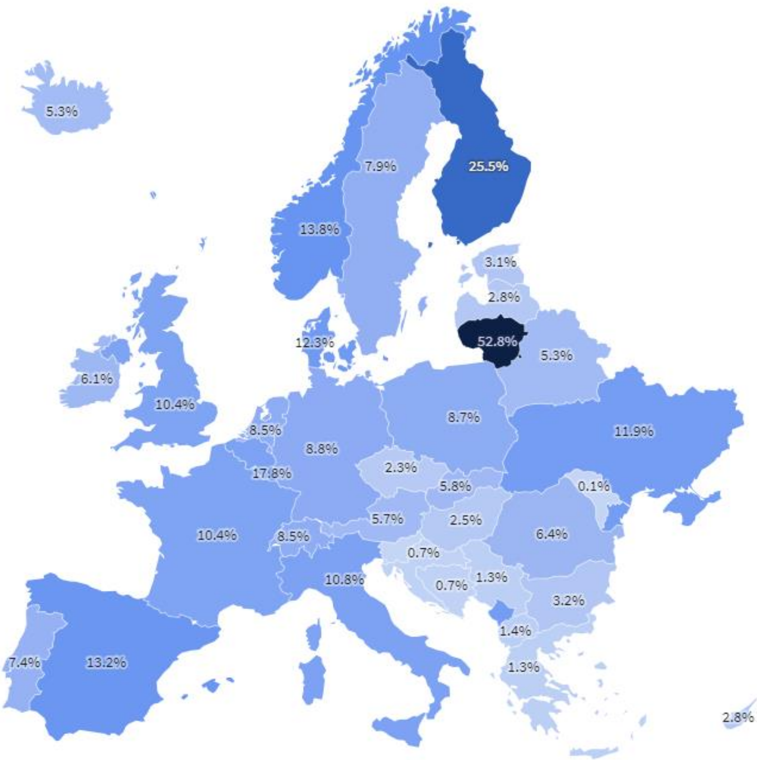

Figure 3: VC amount invested into female founded startups as % of country total (2019 - 2023) - data from Dealroom.com.

The volume and value of investment in women-founded startups varies significantly across countries. Contrary to expectations, more developed ecosystems such as those in the UK and the Nordics do not exhibit a significant advantage in delivering financial support for female founders.

According to a 2018 study by the Boston Consulting Group titled "Why women-owned startups are a better bet," female founders and co-founders typically received less than half of the funding received by male entrepreneurs, but over a span of five years, female-led startups performed better overall and generated 10% more in cumulative revenue.

MANAGEMENT - Human Capital and Skills

The Management factor focuses on the skills, knowledge, and networks that women possess or need to develop for effective business leadership.

Although education has long been considered essential to establishing a company, especially to close the gender gap in entrepreneurship, new evidence indicates that education is typically associated with employment. Gender parity in entrepreneurship is, however, only marginally impacted by equal educational achievement between males and females. This implies that the advantages of educational parity go beyond entrepreneurship and are not sufficient on their own to close the gender gap in this area. On the contrary, perceptions of ability have a stronger direct correlation with bridging the gender gap in entrepreneurship. This could be reflected in the fact that practical entrepreneurial activities do not only require theoretical knowledge, but business knowledge and know-how, even though there may be other cases when they need it, for example, for networking purposes. Alternatively, a more efficient way to support female entrepreneurs is through entrepreneurial education. Gender norms play a crucial role in shaping management methodologies and values, which can impact the relationship between gender representation and sustainability in entrepreneurial ventures. Given that these norms are context-specific, the degree of gender equality within entrepreneurial ecosystems affects female entrepreneurship and the sustainability of startups led by women (D’Ambrosio A., Colombelli A., Ravetti C., 2024).

Although formal education doesn’t bridge the gap in entrepreneurship, the level of education is found to be significant in the case of emerging startups both with a prevalence of female entrepreneurs and of young entrepreneurs.

From an academic perspective, it can be noted that when female entrepreneurs launch their own firms, they frequently have greater levels of education than their male counterparts. According to this trend, businesswomen are more likely to be "over-educated" with the effort to compensate for any potential gender biases or difficulties they may face in the entrepreneurial environment.
Along with the traditional notion of gender roles, the absence of female role models and success stories, women's leadership styles are further impacted.

MARKET - Conditions and Opportunities

Startups are often formed in response to market opportunities. A region's economic structure, demand for innovation, and availability of unmet needs all contribute to entrepreneurial activity. For example, startups often emerge to fill gaps in the market or respond to new consumer demands created by technological advances or social changes. The market dimension examines how entrepreneurs interact with and penetrate markets, including how gender stereotypes can influence women's success in different markets.
With regards to female entrepreneurs, it is reported that starting a business enables women to overcome the "glass ceiling effect" in addition to taking on risks and evaluating their working approach. In a 1978 speech to the Women's Action Alliance, Marilyn Loden first introduced the expression of "glass ceiling effect", describing it as an invisible barrier that prevents certain groups of people, typically women and minorities, from advancing in their careers and reaching high-level positions within organisations or industries. This barrier may present itself in a variety of ways, including selective recruiting procedures, restricted access to programs for training and development, lack of mentorship opportunities, and biased selections on promotions. Thus, the term "glass ceiling" refers to the additional barrier that women face while trying to get into senior positions inside a company.
A study in Organisation for Economic Cooperation and Development (OECD) countries on cross-country labour market conditions and dynamics has been conducted to analyse their influence on female entrepreneurship.
Based on the findings, it's important to highlight that a high rate of female labour force participation along with a smaller gender wage gap improves the conditions for female entrepreneurship. A high percentage of female labour force participation combined with a low representation of women in positions of power represents another path to greater

female entrepreneurship. This situation frequently inspires women to launch their own companies. Furthermore, a country's robust economic, financial, and development circumstances—also denoted by a high-country risk ranking—further encourage female entrepreneurship by boosting investor confidence and resolving funding obstacles that women frequently face.

MACRO - MESO ENVIRONMENT

Support from the environment is one of the major aspects impacting one's decision to start a business. The startup foundation rate of a region is significantly influenced by both macro and meso environmental factors. At the macro level, broader economic conditions, such as national-level policies, economic stability, laws, access to capital, and regulatory frameworks, create the foundational landscape for entrepreneurial activity.

Based on the OECD's 2021 study "Entrepreneurship Policies through a Gender Lens," women's entrepreneurship policy is still a work in progress rather than a finished product. This is because there is evidence that institutional and market barriers, such as societal stereotypes that discourage women from starting businesses and market failures that make it more difficult to access resources, are preventing women from pursuing entrepreneurship. But as the gender gap has gradually decreased over the past 24 years, interest in innovative businesses managed by women has grown due to their potential to create long-term economic growth through their beneficial impacts on employment and wealth.

The local ecosystem, which includes regional industry-specific elements and the existence of institutions that support startups like incubators and accelerators, is critical in forming the startup culture at the meso level. Together with support programs and funding options, family support has been identified in more recent literature as an essential success factor for a startup's starting phase. However, with reference to the presence of incubators in a given area, it does not always appear as fundamental in explaining local birth rates of innovative startups with a prevalence of female founders. This result could be explained by the fact that female entrepreneurs receive less benefit from this form of support because they do not gain a perceived advantage from it, making this factor irrelevant.

MOTHERHOOD

Despite progress towards equity, society still attributes women the dual role of primary caregivers and human resources in the labour market.

Gender roles in society can have negative influences on the scale and nature of women's entrepreneurship. Tax and family laws continue to reinforce traditional gender roles in many OECD nations. Income tax policies that reward households with a single income earner may discourage women from pursuing entrepreneurship. Furthermore, despite changes in family policy that encourage women to enter the workforce, there is still a bias in favour of employment over entrepreneurship. Parental leave and childcare regulations serve as examples of this, as they can significantly affect the feasibility of entrepreneurship for many women.

In recent years, there has been a growing governmental concern about whether paid labour and caring responsibilities can coexist, given that employee requests for flexibility can include changes to the time and location of their place of employment in addition to hours worked. Men are more likely to prioritise standard economic factors when making decisions about self-employment, according to research on the subject. In contrast, women are more likely to prioritise social factors like work-life balance, flexible work schedules, parenthood, childcare responsibilities, and self-esteem concerns.

Beside the social expectations, women experience inequality at home since they bear a greater share of the family's duties. This increases their dependence on financial resources and limits their interpersonal power. This propensity has negative effects on the founding and survival rate of women-led businesses: in 2022, women identified loss of profitability (24.2%), pandemic-related reasons (16.4%), and family and personal difficulties (18%) as the main reasons for leaving a company. When combined, these three explanations make up almost two thirds of the reasons given by women for leaving. About 43% more women than men—one in five—reported leaving their company for family-related reasons. All of these factors reinforce the negative perceptions that prevent even the most privileged women from entrepreneurship and restrict their ability to obtain funding, legitimacy, and other essential resources for the expansion and success of their business ventures. Policies that enhance the availability of household support services, such as eldercare, after-school programs, affordable childcare, along with measures that promote workplace flexibility and make self-employment as attractive as traditional employment, can play a crucial role in reducing the constraints women face due to gender differences in household responsibilities. Following this perspective, researchers have examined the effects of female low-skilled immigrants' labour on the gender wage gap and labour participation gap, only to discover that it actually helps high-skilled native women advance in their careers by enabling them to work in more hours per week.

2.3 Low-skilled Migration and Economic Dynamics

For the purpose of this thesis, it is of high relevance the literature explaining the role of other women in enabling and sustaining the careers of female entrepreneurs.

As highlighted in paragraph 2.2.4, female low-skilled immigrants are of particular interest as they are increasingly contributing to the household responsibilities, such as childcare, eldercare, and domestic work. This support effectively mitigates the burden of these duties mostly attributed to women, thereby offering native women a significant time-cost opportunity. This dynamic not only underscores the interdependence between labour markets and migration but also demonstrates the critical, albeit often overlooked, role that immigrant women play in facilitating the career advancement of native women.

However, before investigating the effects and previous studies on the interaction between low-skilled female immigrants and high-skilled native women, it is necessary to first examine the broader role of low-skilled immigrants within the economy and labour market of a country.

2.3.1 Impact on the Local Labour Market

The Migration Research Hub by the International Migration Research Network (IMISCOE), defines low-skilled migration as “the movement of persons holding jobs that do not necessarily require a high level of education or extensive experience.”

Data from Employment and Social Developments in Europe (ESDE) report that, in 2021, in the countries of the EU the share of migrants was higher in sectors characterised by persistent labour shortages such as domestic services, agriculture, and low-wage industries. Their participation in these sectors strengthens the contribution in sustaining economic performance. Nevertheless, immigration may significantly alter the labour market conditions of a country. In the short-run and with flexible wages, an inflow of low-skilled immigrants increases the supply of low-skilled labour, reduces the average salary of the affected sectors by lowering prices; and it raises the cost of the low-skilled locals' complementary labour market. The concentration of migrants in the household services sector has different implications than overall migration and thus deserves a separate analysis. Given the gender division of household work, migrant domestic workers mostly substitute for unpaid female labour, and therefore, have the potential to increase native women's labour supply to the market.

On the contrary, low skilled immigration might crowd out other types of policies or institutions designed to support native mothers' labour force participation, thus, arising concerns around policy implications.

During the last decades, researchers have examined the effects of this phenomenon on the labour supply of high-skilled native women especially in the USA (P. Cortes, 2011) and in Spain (L. Farré, L. González, F. Ortega, 2011). While the results agree upon the fact that low-skilled immigrants increase the hours worked of high-skilled women, different outcomes emerge regarding the effect on labour force participation rate, suggesting that it depends on the education of natives and their family responsibilities.

2.3.2 Interactions with Female Entrepreneurs

The unequal gender distribution of tasks within the household imposes a burden on women's labour supply. This is particularly the case in countries where families are important providers of care for children and the elderly, as a consequence of weak public policies on this matter.

In the literature on female labour supply, the focus is on the lack and costs of care services, with little attention given to the role of female immigrants who largely provide those services. Despite this, the interaction of immigrant women with female natives and their provision of domestic services and children and elder's assistance, leads to a socio-economic impact. The outcome is, indeed, the supply of a substitute for time consuming activities deriving from the decrease of the relative prices, with a stronger effect on high-skilled female natives, as they are the ones who have a higher opportunity cost of their time. The findings indicate that reducing the costs associated with working longer hours may enable women to increase their work hours and benefit from higher wages and, consequently, contributes to narrowing the gender gap in both weekly hours worked and wages, particularly in occupations where longer work hours are highly rewarded. Besides the existing research regarding the role of low-skilled female immigrants in alleviating domestic responsibilities and supporting the professional advancement of high-skilled native women, there is currently no substantial academic evidence demonstrating a direct positive impact of such labour on the foundation rate of female-led innovative startups in a given region. While the reduction in household burdens might theoretically create conditions conducive to entrepreneurial activity, the specific link between domestic assistance provided by low-skilled female immigrants and the emergence of female-led innovative businesses is not clearly defined yet.

2.4 Overview on the Italian Scenario

The Italian startup ecosystem has notably evolved during the past decade, characterised by significant positive trends and a 46.1% increase between 2018 and 2022 (Infocamere) in innovative startups officially registered. These developments have been subject of extensive analysis by economists and researchers, who seek to understand how the entrepreneurial environment and governmental policies have contributed to this beneficial outcome.

The North-West region of the country is specifically known for its dynamic economic environment, which has become a fertile ground for innovative startups, in particular for the women-led ones. However, to fully understand the impact and potential of the female-led innovative businesses, it is essential to first examine the broader landscape of the Italian startup ecosystem. This overview will provide the necessary context to comprehend the specific challenges and opportunities faced by female entrepreneurs in this region, and how the presence of female low-skilled immigrants potentially influences their decision to undertake the entrepreneurial career.

2.4.1 Entrepreneurial Ecosystem in Italy

The Annual Reports published by the Ministry of Enterprises and Made in Italy systematically monitors the growth trajectory of innovative startups within the country. These reports evaluate both the number of enterprises registered as innovative startups according to the definition outlined in the Startup Act discussed in paragraph 2.1.1, as well as their comprehensive economic performance.

As illustrated in Figure 4, the total number of innovative startups in Italy experienced consistent annual growth from 2019 until 2022. However, this upward trend stabilised, eventually leading to a decline in the absolute number of innovative businesses recorded in the official register by the second quarter of 2024. More precisely, in 2020 the number of registrations in the innovative startups' section of the commercial register was 11.983, with around +10% compared to 2019. The number increased by +16,3% in 2021, with 13.999 startups. In 2022 the positive trend was confirmed, as mentioned beforehand, with 14.264 innovative startups in total, highlighting a less meaningful increase of +1.4% with respect to 2021. At the end of 2023, the Registro delle Imprese accounted for 13.393 innovative startups with a decrease of 6.11%, while, according to the latest data referred to the second quarter of 2024, the registrations have reached a number of 12.871, highlighting a constant decrease, even though with a smaller rate.

The decline in the number of registered innovative startups could be attributed to the medium-term economic repercussions of the Covid-19 pandemic in the country. Along with the pandemic consequences, this reduction may also be influenced by the fact that some businesses are no longer listed in the official register, potentially due to their inability to meet certain requirements, such as the five-year-old limit from their year of foundation. In this scenario, the businesses are forced, then, to transition in the innovative SMEs register, which is, indeed, recording an increase, especially starting from 2022.

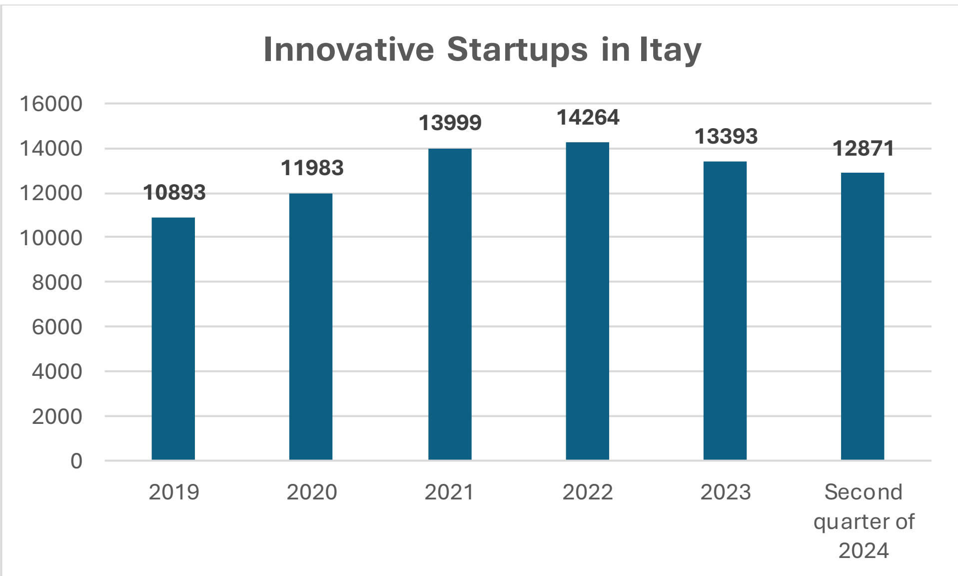

Figure 4: Innovative startups in Italy from 2019 to the second quarter of 2024 (data from Ministero delle Imprese e del Made in Italy)

Regardless of the settlement during the last few years, innovative startups have significantly contributed to the levels of employment of the Italian labour market, registering a +12,5% in 2020. Alongside, the average capitalization of innovative startups is growing, standing at around 89,000 euros per startup following the latest reported status as of second quarter 2024.

In terms of sectoral distribution, the most recent data show that the majority of innovative startups, 78.24%, provide services for businesses, 13.80% are engaged in manufacturing, and 2.77% are involved in commerce.

2.4.2 Innovative Startups in North-Western Italy

An analysis of the geographical distribution reveals that the North-West area of Italy is usually the most flourishing, with the region of Lombardy maintaining its leading position as the region with the highest concentration of innovative startups and numbering 3438, which accounts for 26.71% of the national total. This is followed by Lazio, with 1534 (11.92%) startups, and Campania, which continues its growth trend with 1491 (11.58%) startups. Emilia-Romagna ranks fourth with 898 startups (6.98%), and Veneto follows with 814 startups (6.32%). At the back of the ranking, Basilicata has 119 startups (0.92%), Molise has 84 startups (0.65%), and Valle d'Aosta has only 17 innovative startups, comprising just 0.13% of the national total.

Region	Innovative Startups (data at the second quarter of 2024)	% over total Innovative Startups in Italy
Lombardia	3438	26.71
Piemonte	707	5.49
Liguria	242	1.88
Valle D'Aosta	17	0.13
Total North-West	**4404**	**34.22**
Veneto	814	6.32
Trentino-Alto-Adige	249	1.93
Friuli-Venezia-Giulia	228	1.77
Emilia-Romagna	898	6.98

Total North-East	**2189**	**17.01**
Toscana	565	4.39
Umbria	184	1.43
Marche	293	2.28
Lazio	1534	11.92
Total Centre	**2576**	**20.01**
Abruzzo	309	2.40
Molise	84	0.65
Campania	1491	11.58
Puglia	573	4.45
Sardegna	167	1.30
Basilicata	119	0.92
Sicilia	710	5.52
Calabria	249	1.93
Total South & Islands	**3702**	**28.76**

Table 1: Innovative startups in Italy in the second quarter of 2024, divided per region and geographical area (data from Ministero delle Imprese e del Made in Italy).

Milan continues to lead as the province with the highest number of innovative startups in Italy, with 2,482 startups by the end of the second quarter of 2024, representing 19.28% of the national total. Rome ranks second, being the only other province with over one thousand startups, precisely 1,384 startups, being 10.75% of the total. The other major provinces are significantly behind, with 824 startups (6.40%) in Naples, 511 (3.97%) startups in Turin, and 304 startups (2.36%) in Bari. Conversely, the lowest figures are recorded in Verbano-Cusio-Ossola, followed by Vercelli, with only 2 and 3 innovative startups, respectively.

When analysing the Italian landscape and comparing its cities to other regions within the European ecosystem, Milan ranks 15th in terms of city popularity - % of founders naming the city as a possible startup location for a hypothetical startup. In particular, Milan holds the 6th position for university entrepreneurs, with 1385 startup founders originating from local universities, and has the 2nd largest per capita pool of STEM students. Turin, the main city of Piedmont region, also features in the European rankings, establishing itself as a successful tech hub. Turin's startups raised $70.8 million in 60 funding rounds in 2023, making it the second-largest city in Italy in terms of investment volume.

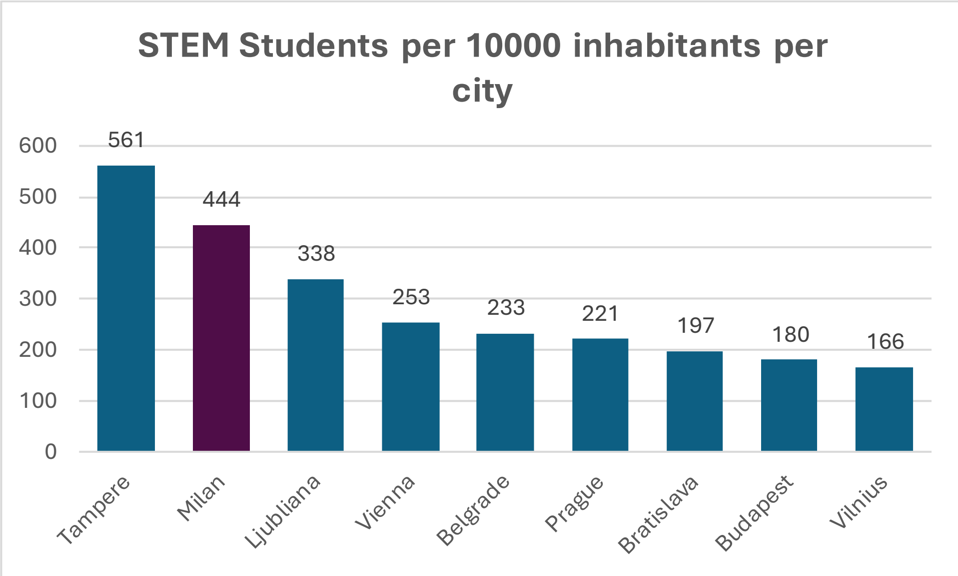

Figure 5: STEM students per 10000 inhabitants per city (data from Startup Heatmap Europe 2023).

It is clear from the data at the Italian territory level that, despite the overall success of the North-Western Italian regions in creating a favourable ecosystem and an evolving environment where to establish innovative startups, the area is very diverse when analysing the provinces separately.

These findings are situated within the broader context of Italy's evolving startup ecosystem, which, despite its relatively recent emergence, has shown remarkable growth. Italy's rich history of entrepreneurship, coupled with its strategic location in Europe, has fostered a strong foundation for startup culture, while always supporting innovation and the risk-taking entrepreneurial spirit. The ecosystem is ranked 15th in Western Europe and 30th globally, with over 40 cities listed among the top 1000 startup hubs. However, the Italian startup ecosystem still faces significant challenges, including the need for a more favourable regulatory environment, improved tax policies, and greater involvement from the private sector. Additionally, while early-stage investment has increased, access to late-stage funding remains a critical obstacle.

Addressing these challenges is essential for sustaining and accelerating the growth of Italy's startup ecosystem, as shown by a study published in 2022 and conducted to examine the broader impact of the Italian Startup Act enforced in 2018. The analysis, based on a sample of innovative startups founded between 2015 and 2018 in 105 Italian provinces, so before and after the reform, indicates that reducing growth barriers has been particularly effective in encouraging individuals with industry-specific, managerial, and entrepreneurial experience to embark on new ventures.

Besides the industrial policies' incentives, the local density of universities and research centres, and the education level of the local population, are particularly impactful in stimulating entrepreneurial activity among highly skilled individuals and in influencing the birth rate of innovative startups.

2.4.3 Female Founders

The presence of female founders in the entrepreneurial panorama is a good indicator of a more inclusive environment, as diversity among founders is a critical determinant for openness to new occasions.

With reference to the composition of the innovative startups with a predominance of women - the startups in which the majority of ownership and administrative positions are held by women, as defined in section 2.1.3 - account for 1835, 14.26% of the total, in 2024. On the contrary, there are 5786 innovative startups in which at least one woman is on the board (innovative startups with female presence).

Within the North-Western area of Italy, which serves as a significant hub for entrepreneurial activity, Milano and Torino stand out as leading cities, not only in terms of the overall number of innovative startups as highlighted beforehand, but also for the proportion of businesses led by female entrepreneurs. Milano, as the primary economic centre of the region, demonstrates a relatively higher percentage of female founders, 17.6%, compared to the European average that stands at 12.98%. As shown in Figure 6, the city is in 6th position in terms of female founders across large European cities, according to the report Startup Heatmap Europe 2023. Turin, while smaller in scale, also shows a noteworthy presence of female-led startups with a percentage of 12.96%, driven in part by the city's strong emphasis on technological innovation and its active participation in startup accelerator programs.

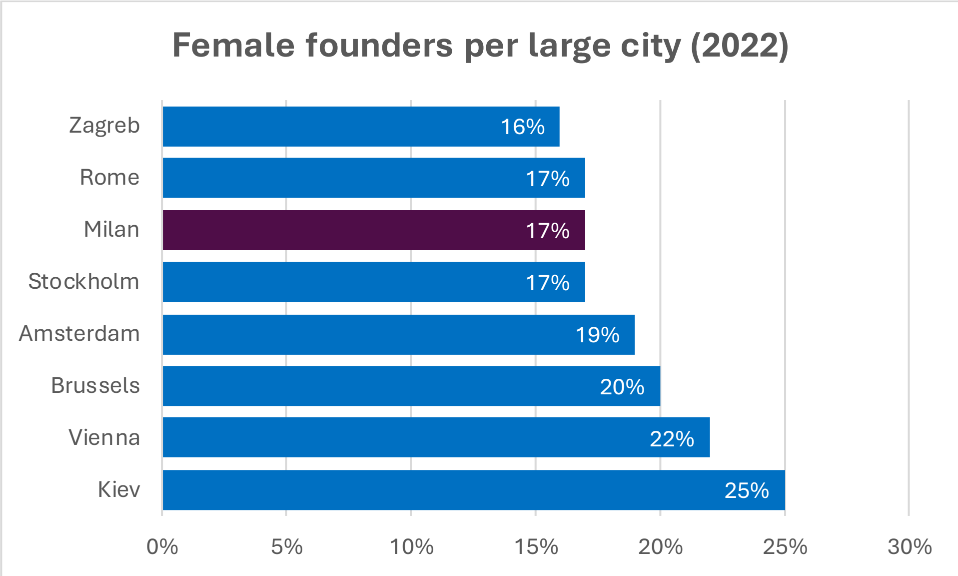

Figure 6: Female founders per large city in 2022 (data from Startup Heatmap Europe 2023).

A quantitative study, based on a sample of over 4,600 Italian innovative startups from the database AIDA, was conducted in 2022 to explore the relationship between gender and innovation intensity of startups, by analysing sales growth from 2015 to 2019. The findings reveal that female-led startups investing more time and resources in innovation experience higher growth rates than their male-led counterparts, challenging the traditional "Female Underperformance Hypothesis." This evidence contributes to the growing literature that highlights the potential of female entrepreneurship, especially in the context of economic recovery following the Covid-19 pandemic.

Female-led companies align with the European average, but significant differences persist in the innovative startup sector. Despite a decline in traditional female-led businesses, the number of female-founded innovative startups in Italy experienced a significant increase between 2018 and 2022, particularly in high-value sectors such as artificial intelligence, edutech, and fintech. This is also reflected in the rise of the so-called research-preneurs, female entrepreneurs from the fields of research translating academic skills into successful businesses. However, challenges such as limited access to specific training and difficulties in securing equitable funding, continue to challenge the full potential of female entrepreneurship in Italy.

In Italy, women who aspire to pursue entrepreneurial careers often encounter significant barriers due to the traditional societal roles attributed to them, especially in managing household activities. Given this context, it is valuable to examine the interplay between female startup founders and the contributions of low-skilled immigrant women who are employed in domestic work. This approach is especially relevant in Italy, where a significant number of immigrants possess low levels of education. Understanding how the support provided by these workers influences the entrepreneurial activities of female founders could offer significant insights about the broader dynamics of gender, labour, and entrepreneurship in the Italian context.

2.4.4 Immigration and Labour Market Dynamics in Italy

Immigration, particularly low-skilled immigration, plays a significant role in shaping the socio-economic landscape of Italy. Over the past decades, indeed, Italy has experienced a considerable influx of low-skilled immigrants, driven by the demand for labour in sectors such as agriculture, construction, and domestic services. Immigration in Italy is largely low-skilled, with a percentage over the total population that is among the highest in Europe. The significance of low-skilled immigration to the nation can be comprehended by taking into account that low-skilled labour is heavily utilised in traditional manufacturing goods, and that the country's ageing population is demanding more personal and household services, which are traditionally provided by low-skilled workers. Additionally, immigrants from South-Eastern Europe and South America, who are often less educated, typically migrate toward Italy.

The North-West of Italy, characterised by its industrial and economic vitality, has attracted a considerable portion of this immigrant population, who often fill essential roles that are less appealing to the native workforce. This region has become a hub for immigrant labour, contributing to the region's economic growth while also presenting challenges related to labour market segmentation and access to social services.

2.4.5 The Role of Female Immigrants in Incentivising Female Entrepreneurship

This thesis investigates the impact of the inflow of female immigrants specialising in household production on the foundation rate of innovative startups led by women in the North-West region of Italy. Thus, the analysis focuses on the causal effect of increased availability of immigrant household services on the work patterns and on the decision to follow an entrepreneurial career of native Italian women.

The findings of the recent studies examined in the paragraph 2.3.2, reveal that a higher presence of immigrant workers providing domestic services leads to an increase in the time native women spend at work, without significantly affecting their overall labour force participation. This effect is particularly pronounced among highly skilled women, whose opportunity cost of time is higher. The evidence suggests that this impact primarily operates through the substitution of household labour rather than complementarities in the production sector, demonstrating that this immigration-driven substitution operates as a private alternative to publicly provided welfare services, raising concerns about the fairness and sustainability of this type of welfare model.

Italy presents a unique case for this analysis, given its increase in low-skilled immigration during the last few decades. By the end of 2008, the country had approximately 3.9 million resident foreigners, comprising 6.5% of the population. In the meantime, the foreign population has increased, with data from ISTAT reporting a number of approximately 5 million foreign nationals living in Italy, representing 8.7% of the total number of residents as of 1 January 2023, which increased by +2.2% compared to the previous year. Notably, the proportion of immigrants who are low-skilled, and particularly those employed in domestic services, is significantly higher in Italy than in other European countries. This situation is further enhanced by the peculiarities of the Italian labour market and social structures, where female labour force participation is influenced by cultural factors and the inadequacy of child and elder care welfare services. The mismatch between employment and childcare availability in Italy amplifies these challenges, as public sector childcare services offer limited hours, which could be appropriate for a non-working or part-time employed parent, and private alternatives are costly and less accessible. Furthermore, in Italy, the price of public childcare is very heterogeneous across regions, as the price is decided by the municipalities and the structures and availability of subsidies is established by the local authorities. Consequently, Italian women face significant constraints in balancing labour market participation with family responsibilities, particularly when compared to their European counterparts.

In a region like North-Western Italy, which includes economic hubs such as the previously mentioned, Milan and Turin, this support could be instrumental in empowering women to transition from employment to entrepreneurship. The ability to outsource household and caregiving tasks may reduce the opportunity cost associated with time-intensive activities such as starting and managing a business. Consequently, the availability of such domestic support could serve as an incentive for female entrepreneurship, contributing to the growth of innovative startups in this area.

3. Empirical Method

Aiming to evaluate the relation between the presence of female immigrants and the foundation rate of female-led innovative startups in the North-Western regions of Italy, this chapter outlines a theoretical overview on the empirical methods employed in this research. To investigate any causal connection between the components under investigation, a statistical model was specifically developed by implementing the linear regression model with both simple and multiple regressors.

3.1 Linear Regression with a Single Regressor

This paragraph introduces the basics in simple linear regression, also known as linear regression with a single regressor, which is a statistical method used to describe the relationship between a variable of interest Y, and an independent, or explanatory, variable X. The relationship between the two variables in the simple linear regression model is therefore represented by a straight line and can be quantitatively expressed with the following formula:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \ ,$$

where

- i is the index running over the n observations, i = 1,2,…,n.
- Y_i is the dependent variable, or regressand.
- X_i is the independent variable, or regressor.
- $\beta_0 + \beta_1 X_i$ is the population regression line, or population regression function.
- β_0 is the intercept of the population regression line.
- β_1 is the slope of the population regression line.
- u_i is the error term.

The intercept $\beta 0$ and the slope $\beta 1$ are the coefficients, or parameters, of the population regression line. In particular, the intercept $\beta 0$ is the value of the population regression line when X=0, which could have a meaningful economic interpretation depending on the application of the model. On the contrary, $\beta 1$ is the change in Y associated with one unit change in X,

$$\beta_1 = \frac{\Delta Y}{\Delta X}$$

Because a model is never able to fully explain a variable of interest Y, a linear regression model always extended by the additive term ui capturing all influences that the model is not accounting for. This "disturbance" term includes all the differences between the regression line and the actual observed data, which, besides pure randomness, could arise from the fact that not all factors affecting Y are included in the model, or from pure randomness.

$$u_i = \mathrm{Y_i} - (\beta_0 + \beta_1 \mathrm{X_i})$$

3.1.1 Estimation of Coefficients

Empirically, model coefficients $\beta 0$ and $\beta 1$ are typically unknown and need to be estimated using a sample of data. Typically, the method that is mainly used to estimate the parameters of the linear regression line is the ordinary least squares - OLS. The objective of the OLS estimator is to select the regression coefficients so that the estimated regression line is as close as possible to the observed data. The method consists in minimizing the sum of squared differences, hence "least squares", between the observed sample values and the fitted values from the model, thus the squared mistakes made in predicting Y given X.

Given the estimators of β_0 and β_1, respectively b_0 and b_1, and given that the regression line that predicts Y as a function of X is $b_0 + b_1X$, the error made when evaluating the value of Yi for the i^{th} observation is

$$u_i = \mathrm{Y_i} - (\mathrm{b_0} + \mathrm{b_1 X_i})$$

Thus, the sum of squared estimation differences can be expressed as

$$\sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$$

The OLS estimators of the intercept $\beta 0$ and the slope $\beta 1$ in the simple linear regression model are calculated by using the formulas reported below.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Xi - \bar{X})(Yi - \bar{Y})}{\sum_{i=1}^{n}(Xi - \bar{X})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X},$$

where:

- $\bar{Y}$ is the least estimator of the population mean E(Y);
- $\bar{X}$ is the least estimator of the population mean E(X).

The estimated intercept $\hat{\beta}_0$, slope $\hat{\beta}_1$ and residuals $\hat{u}_i$ are computed from a sample of n observations of X_i and $Y_i = 1,\ldots,n$. Therefore, the OLS regression line, also known as sample regression line, is given by:

$$\hat{\beta}_0 + \hat{\beta}_1 X_i$$

While the value of Y_i and u_i predicted by OLS regression line are

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

$$\hat{u}_i = Y_i - \hat{Y}_i$$

3.1.2 Measure of Fit

By measuring the fit of estimated linear regression model, it is possible to determine if the observations are closely grouped around the regression line and how well the model describes the data. Both the regression R^2, or coefficient of determination, and the standard error of the regression SER measure how well the OLS regression line fits the data.

The Coefficient of Determination R^2

The regression R-squared is defined as the fraction of the sample variance of Y_i that is explained by X_i. Mathematically, the R^2 is expressed by the ratio of the explained sum of squares to the total sum of squares, where:

- <u>the explained sum of squares ESS</u> is the sum of squared deviations of the predicted values of Y_i, from the sample average $\bar{Y}$ of the *n* observations

$$ESS = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$

- and <u>the total sum of squares TSS</u> is the sum of squared deviations of the Y_i from their sample average $\bar{Y}$ of the *n* observations

$$TSS = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

Thus, the formula to calculate R-squared is the following one

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{ESS}{TSS}$$

Alternatively, given that TSS = ESS + SSR, where SSR is the sum of squared residuals

$SSR = \sum_{i=1}^{n} \hat{u}_i{}^2$, the coefficient of determination R^2 can be calculated as reported below

$$R^2 = 1 - \frac{SSR}{TSS}$$

with $0 \leq R^2 \leq 1$.

In particular, if R^2 assumes values close to 1, it indicates a perfect fit, meaning that $ESS = TSS$, so the model explains all the variation in the dependent variable. Whereas R^2 with values close to 0 indicates that the model explains none of the variation, meaning that $ESS = 0$, thus $\hat{\beta}_1 = 0$, and the independent variable X does not account for any changes in the dependent variable Y. Lastly, intermediate values of R^2 indicate that the model explains that specific percentage of the variations in the dependent variable Y, given the independent variable X.

The Standard Error of the Regression SER

The Standard Error of the Regression is a measure of the average distance between the observed values and the regression line. In other terms, it is an estimator of the standard deviation of the residuals $\hat{u}_i$, consequently, a smaller standard error indicates a better fit. The formula to compute its value is:

$$SER = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n} \hat{u}_i{}^2}$$

The term $n - 2$ in the ratio is given by the fact that the model loses two degrees of freedom from the n observations when estimating the two parameters $\hat{\beta}_0$ and $\hat{\beta}_1$ of the regression line.

3.2 Linear Regression with Multiple Regressors

Considering that the foundation rate of female-led startups is impacted by other factors beyond the presence of female immigrants in the area, as discussed in the literature review, the analysis of this study was conducted by using a multiple linear regression. Linear regression with multiple regressors, also known as multiple linear regression, is an extension of simple linear regression where the model predicts a dependent variable using multiple input, or independent, variables.

3.2.1 Omitted Variables Bias

In case the regressor is correlated with an unobserved variable that has not been included in the statistical analysis and partially influences the regressand, the OLS estimator will be subject to omitted variable bias. Therefore, the biased effect of the omitted variable is attributed to the observed independent variable, resulting in an inaccurate estimate of the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$, thus of the regression line.

As stated in paragraph 3.1, the error term includes a "disturbance" effect arising from omitted variables that affect the value of X_i in the analysis of Y_i. This leads to a correlation between u_i and X_i, revealing the inconsistency of the OLS estimators.

3.2.2 Multiple Regression Model and Control Variables

The multiple linear regression model allows to estimate the effect on the regressand Y_i when changing one regressor X_{1i}, while holding the other regressors X_{2i}, X_{3i},…, X_{ki} constant. In particular, the general formula describing the relationship between a dependent variable Yi and the multiple independent variables X_{1i}, X_{2i},…, X_{ki} is the following

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_j X_{ki} + u_i$$

where

- i = 1,…, n.
- Y_i is the i^th observation on the dependent variable.
- X_{1i}, X_{2i},…, X_{ki} are the i^th observations on each of the j independent variables.
- u_i is the error term. In multiple linear regression models, this term is homoscedastic if the variance if the variance of the residual errors is constant for all the i^th observations and, therefore, does not depend on X_{1i},…, X_{ki}. Otherwise, the error term is heteroskedastic.

- The coefficient β_1 is the expected change on Y_i resulting from a unit change on X_{1i}, holding constant X_{2i},..., X_{ki}. The same reasoning is applied to the other coefficients the predictors, or independent variables.
- The coefficient β_0 is the expected value of Y when all the predictors are equal to 0. As for the simple linear regression, this value can have an economic or only a mathematical interpretation.

3.2.3 OLS Estimators

In accordance with the simple linear regression, also the multiple linear regression uses OLS estimators to predict the population regression line by estimating the values of the unknown coefficients β_0, $\beta_{1,...,}$ β_k. As in the previous case, the value of Y_i is given by

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_j X_{ki}$$

and the residual error is $\hat{u}_i = Y_i - \hat{Y}_i$.

Thus, the population multiple regression model with OLS estimators is given by the following equation

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_j X_{ki} + \hat{u}_i$$

3.2.4 Measure of Fit

Although a multiple linear regression analysis is usually more accurate than a regression with a simple regressor, it is still necessary to evaluate the fit of the estimated model with the observed data. For this purpose, three measures are typically used: the R-squared and the standard error of the regression, as for the previous scenario, and the adjusted R-squared.

The Coefficient of Determination R^2

The R-squared measures the proportion of variance in the dependent variable that is explained by the model. Mathematically, the R^2 can be computed with the same formulas used in the simple linear regression

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{ESS}{TSS}$$

Or it can be rewritten as 1 minus the ratio of the variance of Y_i not explained by the regressors, and considering that $SSR = \sum_{i=1}^{n} \hat{u}_i{}^2$

$$R^2 = 1 - \frac{SSR}{TSS}$$

The Adjusted R-squared

By increasing the number of variables in the model, R^2 is affected by an increase in its value. Consequently, it leads to the idea that the fit of the model improves and, thus, giving a deviated estimate of how well the model fits the observed data.

The $\bar{R}^2$ is a modified version of the R^2 that adjusts for the described misleading behaviour of the latter and is computed as below

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}\frac{SSR}{TSS}$$

The formula shows that the adjusted R-squared is 1 minus the fraction of the sample variance of the OLS residuals to the sample variance of Y, with a correction on the degrees of freedom.

Additionally, considering that the fraction $\frac{n-1}{n-k-1}$ is always greater than 1, then $\bar{R}^2 < R^2$ and the two values are equal when the linear regression has only one regressor.

The Standard Error of the Regression SER

The Standard Error of the Regression measures the standard deviation of the error. In this case, the formula to compute it, is:

$$SER = \sqrt{\frac{1}{n-k-1}\sum_{i=1}^{n} \hat{u}_i{}^2} = \sqrt{\frac{SSR}{n-k-1}}$$

The term $n - k - 1$ in the ratio accounts for the bias introduced by the estimate of the k coefficients of the independent variables and the slope, resulting in a total of k + 1 parameters. However, when the sample size n is large, the effect of the degrees of freedom is negligible.

4. Empirical Application

4.1 Data Construction

The empirical application of this thesis aims to construct a linear regression model with multiple regressors to quantitatively analyse the impact of female low-skilled immigration and local childcare services on the foundation of female-led innovative startups in the North-West of Italy. By using data at the level of provinces, the model estimates the influence of these two factors on entrepreneurial activities among native women, especially in innovation-driven sectors, controlling for other demographic and economic variables. First data are collected from four different databases with different timespans, and then grouped together to construct the final database and execute the statistical analysis. Nevertheless, due to data unavailability from databases regarding childcare services and the control variables in terms of local economic indicators, the timespan considered in the final database of the linear regression goes from 2013 to 2022.

The research utilizes provincial-level data, despite the availability of more granular data at the level of the municipality. This choice is grounded in the assumption that the location of the startup and the supply of childcare services, both from public or private organisations and from the support provided by the female foreign population, are not necessarily confined to the same municipality. In many cases, entrepreneurs may seek support in neighbouring areas due to the absence of such services locally.

Given the structure of each database, which contain yearly observations on variables such as the presence of female immigrants, local childcare services, local economic and demographic indicators and the number of female-led innovative startups over multiple years for different geographical areas, they would qualify as panel data.

Panel data, also known as longitudinal data, is a type of dataset that combines both cross sectional data – data on multiple entities for a single period - and time series data - data on a single entity over multiple time periods (J. Stock & M. Watson, 2012). Panel data is particularly valuable in identifying causal relationships because it enables the study of how variables evolve and interact across both perspectives. In the end, the database with panel data is collapsed to examine the number of startups founded in the 25 analysed provinces across the years of observation.

4.1.1 Startups Database

The startups database comprises data regarding startups that are or were registered in the Innovative Startups Register between 2010 and 2022, maintained and updated by Ministero delle Imprese e del Made in Italy. Nevertheless, these data can be considered reliable and significant only starting from 2012, which corresponds to the regulation introduced by the Italian policy makers to enhance the innovative entrepreneurial activity in the nation. Registration is based on a set of criteria established in the official definition of Innovative Startups, as outlined by the Italian authorities and detailed in paragraph 2.1. In case a startup doesn't stop its business after 5 years of activity or doesn't meet anymore one of the other requirements before the fifth year, it needs to move from the register of startups to the one of enterprises. To have a broader view on the financial and commercial status of these startups, the dataset of the Innovative Startups Register was linked to the one of AIDA, which offers additional variables essential to the statistical analysis.

The resulting dataset is made of 28 variables listed below, that describe the identity and the economic and legal features of the startups at their year of foundation and during the following years up to their registration at the register of Innovative Startups or the one of enterprises.

- **id** is the number that univocally identifies a startup, with its maximum value being an indicator of the total number of startups observed to conduct the study.
- **anno_osservazione** refers to the year of observation of the other variables of the database.
- **anno_costituzione** reports the year of foundation of the startup.
- **comune** is the name of the municipality in which the startup has its headquarters and, therefore, in which it is legally recorded according to the register.
- **codice_istat_comune**: is the unique code provided by ISTAT to identify the municipality.
- **provincia** is the corresponding province to which the municipality belongs.
- **codice_istat_provincia** is the unique code provided by ISTAT to identify the province.

- **regione** identifies the region where the province is located. Precisely, in this case it only considers the North-Western Italian regions – Liguria, Lombardia, Piemonte and Valle D'Aosta - for the purpose of the study.
- **stato_giuridico** denotes the legal status of the startup at the time the study was conducted. Specifically, it either reports "attiva" for the startups or companies whose business is still active, or "cessata (in liquidazione)" for the ones that are no longer active and stopped their business operations.
- **ateco_2007_** is a code provided by ISTAT to classify the economic activities of an enterprise according to the industry where it operates.
- **dipendenti** is the variable referring to the number of employees that work for the startup at the year of observation. Its value can range from zero, in case of individual enterprises.
- **tot_immob_immateriali** reports the thousand amounts of total intangible assets at the observed year, in a string type of data.
- **tot_immob_materiali** reports the thousands of total tangible assets at the observed year, in a string type of data.
- **tot_attivo** reports the value in thousands of total assets of the startup at the observed year.
- **tot_patrimonio_netto** indicates the thousand stockholders' equity amount of the startup at the observed year.
- **tot_debiti** indicates the total thousand amount of debt of the startup at the year of observation.
- **capitale_sociale** shows the total equity in thousands recorded in the balance sheet of the startup at the observed year.
- **ricavi** indicates the value in thousands total turnover generated by the business and recorded in the balance sheet at the year of observation.
- **risultato_operativo** indicates the value in thousands of the operating income generated by the business and recorded in the balance sheet at the year of observation.
- **valore_aggiunto** refers to the added value in thousands generated by the startup at the end of the observed year.

- **utile** records in thousands the income of the business at the end of the observed year.
- **ebitda** stands for Earnings Before Interest, Taxes, Depreciation and Amortization. The variable records the ebitda of the business at the year of observation.
- **diritti_brevetto_industriale** records the number of patents owned by the startup at the observed year.
- **data_iscrizione_registro_startups** is the date (dd/mm/yyyy) of registration to the Ministry's Innovative Startups Register. Normally the year of this date should be equal to the one recorded in anno_costituzione. However, there are cases where this does not happen due to bureaucratic timelines.
- **data_uscita_sezione_startup** is a string variable corresponding to the date (dd/mm/yyyy) the startup exited the Innovative Startups' register.
- **data_cessazione** is a string variable indicating the date the startup failed and officially stopped the business activity.
- **prevalenza_femminile_complessiva** is a string variable indicating the level of female participation aligned with the percentage of women in the Board of Directors and among shareholders. The values of the variable can be set as below, in accordance with the definition outlined by Infocamere.

String values	Percentage
NO	(% women in BoD + % female shareholders)/2 < 50%
MAGGIORITARIA	50% ≤ (% women in BoD + % female shareholders)/2 < 66%
FORTE	66% ≤ (% women in BoD + % female shareholders)/2 < 100%
ESCLUSIVA	(% women in BoD + % female shareholders)/2 = 100%

Table 2: Description of the variable prevalenza_femminile_complessiva

- **prevalenza_femminile_amministrativa** is a string variable indicating the level of female participation in the Board of Directors. The values are estimated according to Table 3 but considering only the percentage of women in the BoD.

4.1.2 Immigrants Database

The dataset with demographic data about the overall population both with an Italian and a non-Italian citizenship from 2009 to 2022 are obtained by the broader database of ISTAT – the national institute of statistics. Given that the objective of this study is to analyse the impact of immigrant women on supporting native women with household and caregiving responsibilities, it was essential to use data on low-skilled women with a foreign citizenship of working age, assumed to be between 20 and 65, as they are less likely to have completed higher education. The immigrant's database is based on a strong assumption. Specifically, the total female foreign population is used as a proxy for the population employed in domestic and caregiving roles. This assumption, while simplifying, is grounded in the reality that a significant share of the foreign population in Italy comes from economically less developed countries, particularly from Eastern Europe and Latino America. Immigrants from these regions tend to accept employment predominantly in low-skilled sectors, including domestic and caregiving labour. As such, this assumption aligns with the observed labour market dynamics in Italy, making it a realistic basis for the analysis.

The database, composed of 250 observations in rows, was extracted from the ISTAT by filtering foreigners in the population resident in Italy and setting the age range as mentioned above and the gender equal to female. It resulted in a database with the following 4 variables:

- **provincia** is the geographical indicator referring to the name of the province of the observed demographic data. As the study is conducted in the North-West of Italy, 25 different provinces' names from the regions of Liguria, Lombardia, Piemonte and Valle D'Aosta can be found in this variable.
- **anno_osservazione** records the year of observation, going from 2013 to 2022, of the observed data. Thus, for each province there are 10 years of observation of the below variables.
- **donne_immigrate** has the numerical data regarding the number of immigrant women between the age of 20 and 65 in a specific province in the observed year.
- **tot_popolazione** is, instead, referred to the total number of people living in a specific province during the observed year, without filters on gender, age and citizenship.

4.1.3 Local Services of Childcare Database

To conduct the study, data about caregivers – precisely childcare and eldercare - at the level of at least the province of the analysed regions were needed. Nevertheless, with regards to the eldercare services, the national institute for statistics in Italy has only collected data at a regional level about the structures offering this support. Taking into consideration that a regional-level analysis would have resulted in inconsistent and inaccurate outcomes, only data about childcare service providers were collected for the scope of this thesis.

Within the ISTAT database, the category of "socio-educational services for early childhood" was selected in order to provide the necessary data to complete the study. Specifically, this category comprises nurseries, including corporate nurseries, early childhood sections known as "sezioni primavera" and supplementary early childhood services such as play areas, home-based services, and parent-child centres, both public and private.

The first year of observation at disposal for the early childcare services is 2013, while the last one is 2022.

Kindergarten, or preschool, data were not considered and, therefore, not collected in this database because it is a service that almost every parent benefits from in Italy, even though it is not mandatory for children to attend it could be provided by private institutions.

As a result of this scenario, the database of childcare providers is composed of 250 observations in rows and 3 variables in columns, listed below.

- **provincia** is the geographical indicator referring to the name of the province of the observed data. As the area under analysis is the same throughout all the study, this variable can assume the same values as in the database with demographic indicators.
- **anno_osservazione** records the year of observation of the data under analysis starting from 2013 to 2022, because of data unavailability before and after these years.
- **tot_servizi_prima_infanzia** is the total number of structures providing early childhood services, which includes the subcategories mentioned in the earlier description.

4.1.4 Control Variables Database

As noted in Section 2.2, the determinants of entrepreneurship and the establishment of innovative startups are diverse. Consequently, to develop a robust and precise statistical model, it is essential to incorporate demographic and economic indicators of the areas under examination as control variables.

The database is downloaded from ISTAT, as per the immigrant population and local services of childcare databases, and its years of observation on the indicators at the level of the province go from 2012 to 2022. In particular, it includes 275 observations in rows and 9 variables in columns listed below.

- **anno_osservazione** is the year of observation of the below described variables.
- **population** refers to the total population resident in a province. It has the same value as the variable tot_population in the immigrants database and is, thus, deleted before the merging of the databases into the final one.
- **working_population** is the total population that is actively involved in the labour market of a province. When compared to the total population, the indicator helps to comprehend how effectively the population's labour potential is being utilized.
- **unemployment_rate** is the share of the labour force without work and is obtained by computing the ratio between the unemployed population and the total labour force.
- **value_added_per_habitant** refers to the average economic value generated by each individual in a specific province. It is generally computed by dividing the total Gross Value Added (GVA) of the region by its total population.
- **number_firms** is the variable referring to the total number of firms regardless of their legal structure in a given province at the year of observation.
- **area_km2** indicates the area of the observed province expressed in km.
- **provincia** is the geographical indicator referring to the name of the province of the observed data and it can assume the same values as per the other databases.
- **residenti_laureati** is the number of graduated inhabitants in the province in the observed year.

4.1.5 Final Database

To perform the linear regression analysis, it was first necessary to consolidate all the variables from the aforehand described databases into a single dataset. However, before merging the datasets, it was essential to clean the data by removing observations that are irrelevant to the objective of the research and deleting variables that appeared in multiple datasets, such as the total population of a province in the year of observation, or those considered not relevant for the reliability and accuracy of the model, such as the name of the municipality where the startup is registered and its unique code.

With regards to the dependent variable - the foundation rate of female-led innovative startups - it is never mentioned among the already defined variables. Therefore, it was computed starting from the variables of the startup database and being consistent with the definition given in Paragraph 2.2. Specifically, it was calculated as the ratio between the number of women-led innovative startups and the total number of firms in a given province at the year of observation.

The other additional variables constructed both to calculate the regressand and to conduct the analysis are listed below.

- **startup_prev_femminile** is a dummy variable equal to 1 if the variable prevalenza_femminile_complessiva shows values equal to “ESCLUSIVA”, “FORTE” and “MAGGIORITARIA”, and equal to 0 if prevalenza_femminile_complessiva is equal to “NO” or is empty.
- **tot_startup_innovative** is a variable used to count the total innovative startups in the database by initially setting the value equal to 1 for all the observations where the year of observation is equal to the year of foundation of the startup. This variable, together with the previous one, will be then summed up by anno_costitutzione and provincial with the command “collapse” in STATA.
- **female_startup_rate** is the dependent variable obtained for each year of observation – thus, year of foundation – and province, with the following formula

  $$female\_startup\_rate = \frac{startup\_prev\_femminile}{number\_firms}$$

  where the numerator and denominator are the already summed up values.
- **female_startup_rate_1000** is a variant of the above regressand, scaled by 1000. It represents the foundation rate of women-led innovative startups per 1000 firms.

- **female_startup_share** is a variant of the definition of startup foundation rate outlined in Paragraph 2.2. It is later utilized to develop an alternative statistical model, where the emphasis shifts from the size of women-led innovative startups to their gender composition.

After defining the additional variables and integrating all data into the final dataset, a total of 250 observations were obtained, considering that this is based on the analysis of 25 provinces over a 10-year period going from 2013 to 2022.

4.2 Hypothesis Formulation and Model Construction

Building upon the existing literature, particularly the work of Patricia Cortés, this thesis studies how the presence of low-skilled foreign women as a support to local caregiving services reduces the domestic and caregiving responsibilities traditionally sustained by women, focusing especially on the impact that it has on high-skilled native's women decision to launch an innovative startup. The central hypothesis of the research is that a higher proportion of female immigrants and a higher amount of local caregiving services for early childhood positively impact the foundation rate of female-led innovative startups at the provincial level. Nevertheless, this starting point is further explored to investigate potential shifts in the gender composition of innovative startups, beside simply examining the scale of female entrepreneurship relative to the total number of firms in a certain area. To test these hypotheses, a linear regression model with multiple regressors is constructed. The models analyse provincial-level data spanning from 2013 to 2022, incorporating variables that measure the foundation rate of female-led startups, the proportion of foreign female immigrants, and the availability of local care services. The models also account for a range of control variables, including demographic and economic factors, to isolate the impact of immigration and childcare services on female entrepreneurial activities.

The first tested hypothesis is as follows:

Hp1. *"The foundation rate of female-led innovative startups, relative to the total number of firms in each province of the North-West of Italy, increases with a rise in the rate of foreign female population and in the availability of early childhood care services within the same province."*

The statistical model constructed to test the above hypothesis is a linear regression model with multiple regressors, as described below.

$$StartupFoundationRate_{jt} = \beta_0 + \beta_1 FemImmRate_{jt} + \beta_2 TotServices_{jt} + \beta_3 X'_{jt} + \varepsilon_{jt}$$

where:

- **$StartupFoundationRate_{jt}$** is the dependent variable, that is the foundation rate of innovative startups led by women with respect to the overall firms in a province j in the observed year t.
- **$FemImmRate_{jt}$** is one of the independent variables, capturing the proportion of female foreign population with respect to the overall population in a province j in the observed year t.
- **$TotServices_{jt}$** is another independent variable, and represents the total number of available early childhood care service facilities in a province j in the observed year t.
- **X'_{jt}** is a vector including all the control variables, accounting for additional factors related to province j in year t to ensure a more accurate analysis.
- **β_0** is the intercept, representing the expected value of the dependent variable when all independent variables and control variables are held constant.
- **β_1** is the coefficient of the independent variable FemImmRate, indicating its effect on the foundation rate of female-led startups.
- **β_2** is the coefficient of the independent variable TotServices, representing the impact of early childhood care services on the dependent variable.
- **β_3** represents the coefficients of the control variables in vector X'_{jt}.
- **ε_{jt}** is the error term, capturing unobserved factors that may influence the dependent variable but are not included in the model.

For a deeper understanding of the dynamics between foreign women, local support services and the foundation rate of startups led by women this study also investigates the effect of the simultaneous increase the two first subjects.

The hypothesis stated below, indeed, examines both the individual and combined effects of the proportion of female immigrants and the availability of local childcare services on the foundation rate of female-led innovative startups. Specifically, the model following the hypothesis introduces an interaction term to assess whether the impact of one factor depends on the level of another factor.

Hp2. *"The foundation rate of female-led innovative startups, relative to the total number of firms in a province, is positively impacted by both a higher proportion of the foreign female population and a greater availability of early childhood care services. Additionally, the relationship between these two factors is interdependent, such that a combined increase of both variables is more likely to positively affect the foundation rate of female-led startups."*

$$\begin{aligned} StartupFoundationRate_{jt} \\ &= \beta_0 + \beta_1 FemImmRate_{jt} + \beta_2 TotServices_{jt} + \beta_3 FemImmRate_{jt} \\ &* TotServices_{jt} + \beta_4 X'_{jt} + \varepsilon_{jt} \end{aligned}$$

Where the terms correspond to the variables and coefficients of the previously described model, with the exception of the following:

- **$FemImmRate_{jt}$*$TotServices_{jt}$** is the interaction term between **$FemImmRate_{jt}$** and **$TotServices_{jt}$**, capturing how the effect of the female immigrant population on the foundation rate of startups may vary depending on the availability of early childhood care services in the province. This interaction enables the understanding of whether the relationship between female immigrant presence and startup formation is strengthened or weakened by the level of childcare support.
- **β_3** in this case is the coefficient for the interaction term, indicating the direction of the influence of the combined effect of **$FemImmRate_{jt}$** and **$TotServices_{jt}$** on the foundation rate of female-led innovative startups.

The final model aims to further explore the dynamics of female entrepreneurship in relation to male entrepreneurship within the realm of innovative startups. Therefore, it modifies the previously used definition of the dependent variable and tests the following hypothesis.

Hp3. *"The foundation rate of female-led innovative startups, relative to the total number of innovative startups in each province of the North-West of Italy, increases with a rise in the rate of foreign female population and in the availability of early childhood care services within the same province."*

The following linear regression model with multiple regressors used to test this hypothesis, shows similarities with the one tested in Hp1. Indeed, the main and only change is on the different interpretation used to calculate the dependent variable.

$$FemStartupRate_{jt} = \beta_0 + \beta_1 FemImmRate_{jt} + \beta_2 TotServices_{jt} + \beta_3 X'_{jt} + \varepsilon_{jt}$$

Where the independent variables and control variables under analysis are the same as in the first model, and:

- **$FemStartupRate_{jt}$** is the dependent variable, that is, in this case, the foundation rate of innovative startups led by women with respect to the overall innovative startups in a province j in the observed year t.
- **β_0** is the intercept, representing the expected value of the dependent variable when all independent variables and control variables are held constant.
- **β_1** is the coefficient of the independent variable FemImmRate, indicating its effect on the foundation rate of female-led startups.
- **β_2** is the coefficient of the independent variable TotServices, representing the impact of early childhood care services on the dependent variable.
- **β_3** represents the coefficients of the control variables in vector X'_{jt}.
- **ε_{jt}** is the error term, capturing unobserved factors that may influence the dependent variable but are not included in the model.

4.3 Variable Description

This section describes the dependent, independent, and control variables that form the basis of the linear regression models used to test the research hypotheses. It is, indeed, essential to outline the variables incorporated into the study, in order to conduct a comprehensive analysis.

4.3.1 Dependent Variable

The dependent variable in this analysis captures the foundation rate of female-led innovative startups, and it is computed in two distinct ways to reflect different aspects of female entrepreneurship. Conducting the statistical analysis by using both definitions enables a wider understanding of female-led entrepreneurial activities, both in terms of their overall size within the general business landscape and of their relative share within the innovative startup ecosystem. However, it is expected that the first interpretation will show more significant and valuable results with regards to the research objective. On the contrary, the second one is less likely to provide relevant insights due to its limited dimension with respect to the overall business context.

Startup foundation rate and startup foundation rate per 1000 enterprises

First, the variable is calculated as the ratio of female-led innovative startups to the total number of firms for each province of the North-West of Italy in the observed years, providing an understanding of the overall size of female-led innovative businesses within the local business environment. To further refine this measure, the analysis also includes a variant where the foundation rate is multiplied by 1000, yielding the female-led startup foundation rate per 1000 firms. This transformation allows for a more direct comprehension of the quantitative impact that the regressors have on the regressand.

Startup foundation share

The second definition of the dependent variable focuses on the information regarding the gender composition of the innovative startup ecosystem, calculated as the ratio of female-led innovative startups to the total number of innovative startups in each province of the North-West of Italy in the observed years. This measure emphasizes the proportion of innovative startups founded by women, offering insight into gender representation within the innovative sector specifically, rather than the broader business landscape.

4.3.2 Independent Variables

By including the two following independent variables, the model investigates how the presence of female immigrants and the availability of early childhood services shape the conditions for female entrepreneurship in the innovative business environment.

Rate of female immigrants

The first independent variable is the rate of female immigrants, which is calculated as the ratio of foreign women aged 20 to 65 to the total population of each province of the North-Western regions of Italy, without restrictions on gender, nationality, or age. This variable aims to measure the presence and proportion of female immigrants of working age, whose contribution to the local labour market, particularly in caregiving and domestic roles as a support to local services, is hypothesized to influence the opportunity and decision to start a business of native women.

Total support services for early childhood

The second independent variable is the rate of services for early childhood, defined as the ratio of local facilities dedicated to early childhood support to the total population in each province of the North-Western regions of Italy. This regressor acts as an equivalent of the availability of childcare services, which is expected to reduce the caregiving workload sustained by women and enable their possibilities to embark on an entrepreneurial career.

4.3.3 Control Variables

As discussed in Paragraph 4.2, the models incorporate demographic and economic factors useful to account for the level of development and wealth in each province, as these might have a significant influence on the entrepreneurial opportunities and, consequently, on the startup foundation rate, that is not included in the independent variables. Even though they are presented as a vector in the model description, the control variables described below are individually added alongside the independent variables when performing the analysis.

Rate of graduated residents

The rate of graduated residents in the observed year is included among the control variables as it is a clear indicator of the highly educated human capital available present in a province, which is a key enabling factor in the foundation of innovative startups. Graduated residents

are, indeed, more likely to have the technical skills and managerial know-how needed to establish and scale-up innovative startups.
The rate is computed as the ratio between the number of graduated residents and the overall population of the province during the observed year.

Unemployment rate

The unemployment rate is the share of the labour force without work, calculated as the ratio of unemployed residents of an area to the total labour force, and an important indicator of economic and social well-being (OECD, 2013).
High unemployment, indeed, lead to lower disposable income among the population, making it more difficult for new businesses to thrive. On the contrary, in regions with low unemployment, financial institutions may be more willing to lend given that signals of strong economic performance reduce the perceived risks of business failure.
Given the economic implications of the unemployment rate, it is included as a control variable in the statistical analysis to account for its potential effects on entrepreneurial activity. By doing so, the model adjusts for variations in local labour market conditions that could influence both necessity and opportunity-driven entrepreneurship. For instance, in regions with high unemployment, self-employment might increase as individuals, including women, turn to entrepreneurship out of necessity rather than opportunity. Despite the increase in the foundation rate of female-led startups, these businesses could be driven more by survival than by innovation, potentially limiting their long-term success (Amoros et al., 2019).

Value added per inhabitant

The value added per inhabitant is given by the rate of gross value added of a province to its total population, where the gross value added is measured as the difference between total production and intermediate consumption - the value of goods and services used in the production process - and it reflects the value created by the production of goods and services after subtracting the cost of inputs.
When divided by the total population, it measures how much economic value is created on average by each inhabitant and provides insights about the potential standard of living in the region. Higher value added per capita generally suggests better living conditions, as it implies greater wealth generation, potentially leading to higher wages and more resources

available for public services for the local authorities to invest in, such as infrastructures, healthcare, education and social services.

Governments and policymakers use this metric to evaluate economic development and growth potential. Therefore, it is expected that a rising value added per inhabitant typically indicates economic progress and greater opportunities for innovation and business growth.

Area in square kilometres

Among the control variables, the area of the provinces expressed in square kilometres needs to be included, as it is related to population density, which can significantly affect the entrepreneurial context. The size of a province can also influence the competitive landscape for startups. In larger areas, startups may indeed face higher levels of competition compared to smaller provinces.

Considering this control variable enhances the robustness of the findings, allowing for a clearer interpretation of the factors that drive female-led innovative entrepreneurship in the provincial contexts.

Total population

Even though the total population is always used to calculate the rates of the other independent and control variables, it is relevant to employ its levelled value among the other control variables as it might account for additional demographic insights. The population size, in fact, reflects the scale of the potential market, availability of a diverse and skilled workforce, and access to financial resources, all of which play a crucial role in determining the sustainability of the foundation of new businesses.

4.4 Descriptive Statistics

Once the database that is used to test the hypothesis is set, it is significantly useful to execute the descriptive statistics analysis before diving into the results of the linear regression analysis. The descriptive statistics analysed in this paragraph might, indeed, anticipate the final output by showing trends in the distribution of variables and connections within variables.

anno_osservazione

The variable anno_osservazione refers to the years of observation going from 2013 to 2022, that is the timeline considered for the analysis.

anno_osservazione	Freq.	Percent	Cum.
2013	25	10.00	10.00
2014	25	10.00	20.00
2015	25	10.00	30.00
2016	25	10.00	40.00
2017	25	10.00	50.00
2018	25	10.00	60.00
2019	25	10.00	70.00
2020	25	10.00	80.00
2021	25	10.00	90.00
2022	25	10.00	100.00
Total	250	100.00	

Figure 7: Frequency, percent and cumulative percent of years of observations.

Figure 7 shows the distribution of the variable, where each observed year has an equal frequency of 25 corresponding to the total provinces of the Noth-Western regions of Italy observed for the purposes of the study, contributing to 10% of the total, with a cumulative total reaching 100%.

provincia

This is the variable indicating the geographic area of the regressand and of the other regressors. Therefore, the table of frequency below shows the 25 provinces of the regions of Liguria, Lombardy, Piedmont and Valle D'Aosta, that are observed 10 times, once a year from 2013 to 2022. The total number of observations is 250, accounting for 100% of the data.

provincia	Freq.	Percent	Cum.
Alessandria	10	4.00	4.00
Asti	10	4.00	8.00
Bergamo	10	4.00	12.00
Biella	10	4.00	16.00
Brescia	10	4.00	20.00
Como	10	4.00	24.00
Cremona	10	4.00	28.00
Cuneo	10	4.00	32.00
Genova	10	4.00	36.00
Imperia	10	4.00	40.00
La Spezia	10	4.00	44.00
Lecco	10	4.00	48.00
Lodi	10	4.00	52.00
Mantova	10	4.00	56.00
Milano	10	4.00	60.00
Monza e della Brianza	10	4.00	64.00
Novara	10	4.00	68.00
Pavia	10	4.00	72.00
Savona	10	4.00	76.00
Sondrio	10	4.00	80.00
Torino	10	4.00	84.00
Valle d'Aosta/Vallée d'Aoste	10	4.00	88.00
Varese	10	4.00	92.00
Verbano-Cusio-Ossola	10	4.00	96.00
Vercelli	10	4.00	100.00
Total	250	100.00	

Figure 8: Frequency, percent and cumulative percent of provinces.

startup_prevalenza_femminile

This is the variable indicating the number of innovative startups led by women in a certain province and year and is used in the ratio to compute the dependent variable for all the models.

Variable	Obs	Mean	Std. Dev.	Min	Max
startup_pr~e	250	9.824	30.57417	0	242

Figure 9: Summary statistics of female-led startups' variable.

The statistics indicate that the minimum value of 0 reflects at least one province where no female-led innovative startups were founded in a certain year. In contrast, the maximum value of 242 represents the largest number of innovative startups led by women founded in a single year within a province. The mean of 9.824 suggests that, on average, around 9.8 startups were founded per year in each province, while the standard deviation of 30.57 points to substantial variation in startup activity across provinces and years. Furthermore, the value of the standard deviation and a mean value closer to the min rather than the max suggest a distribution where female-led innovative startups founded every year in each province are only a few and, as reported in Figure 10 below, zero. In contrast, the highest values are expected to be registered in specific provinces characterised by a dynamic entrepreneurial ecosystem.

(sum) startup_prev_femminile	Freq.	Percent	Cum.
0	66	26.40	26.40
1	44	17.60	44.00
2	27	10.80	54.80
3	23	9.20	64.00
4	17	6.80	70.80
5	11	4.40	75.20
6	2	0.80	76.00
7	7	2.80	78.80
8	4	1.60	80.40
9	5	2.00	82.40
10	3	1.20	83.60
11	4	1.60	85.20
12	3	1.20	86.40
13	5	2.00	88.40
14	4	1.60	90.00
15	1	0.40	90.40
17	2	0.80	91.20
18	1	0.40	91.60
19	1	0.40	92.00
20	3	1.20	93.20
21	1	0.40	93.60
22	1	0.40	94.00
29	1	0.40	94.40
31	2	0.80	95.20
33	1	0.40	95.60
42	1	0.40	96.00
54	1	0.40	96.40
67	1	0.40	96.80
107	1	0.40	97.20
108	1	0.40	97.60
126	1	0.40	98.00
169	1	0.40	98.40
172	1	0.40	98.80
177	1	0.40	99.20
220	1	0.40	99.60
242	1	0.40	100.00
Total	250	100.00	

Figure 10: Frequency, percent and cumulative percent of founded innovative female-led startups.

tot_startup_innovative

This variable represents, instead, the total number of innovative startups founded in a single year within a province, without a filter on the gender of the shareholders and individuals of the BoD and is used to obtain the dependent variable for the third model.

Variable	Obs	Mean	Std. Dev.	Min	Max
tot_startu~e	250	35.584	100.8452	0	753

Figure 11: Summary statistics of innovative startups' variable.

In accordance with the previously described variable, the minimum number of innovative startups founded in a year in a certain province is zero. However, in this case, the maximum value is of 753, indicating that male-led startups are more likely to be founded in the North-West of Italy. A mean value of 35.584 – higher than the one in Figure 9 - confirms that, in the analysis of the overall landscape, a larger number of innovative startups is founded. Nevertheless, the mean and the standard deviation of 100.845 support the frequencies in Figure 11 that only a few innovative startups are founded within 2013 and 2022 in the observed provinces, with some exceptions.

(sum) tot_startup _innovative	Freq.	Percent	Cum.
0	10	4.00	4.00
1	26	10.40	14.40
2	15	6.00	20.40
3	16	6.40	26.80
4	14	5.60	32.40
5	16	6.40	38.80
6	11	4.40	43.20
7	11	4.40	47.60
8	15	6.00	53.60
9	7	2.80	56.40
10	5	2.00	58.40
11	7	2.80	61.20
12	6	2.40	63.60
13	5	2.00	65.60
14	3	1.20	66.80
15	3	1.20	68.00
16	3	1.20	69.20
17	4	1.60	70.80
18	2	0.80	71.60
19	1	0.40	72.00
20	2	0.80	72.80
21	2	0.80	73.60
22	3	1.20	74.80
23	1	0.40	75.20
24	2	0.80	76.00
26	2	0.80	76.80
27	1	0.40	77.20
28	2	0.80	78.00
29	2	0.80	78.80
30	2	0.80	79.60
31	2	0.80	80.40
32	1	0.40	80.80
33	1	0.40	81.20
35	4	1.60	82.80
36	2	0.80	83.60
39	2	0.80	84.40
41	3	1.20	85.60
44	1	0.40	86.00
45	1	0.40	86.40
50	1	0.40	86.80
51	2	0.80	87.60
53	1	0.40	88.00
55	2	0.80	88.80
57	1	0.40	89.20
62	1	0.40	89.60
64	2	0.80	90.40
65	1	0.40	90.80
66	1	0.40	91.20
70	1	0.40	91.60
73	2	0.80	92.40
81	1	0.40	92.80
85	1	0.40	93.20
88	1	0.40	93.60
95	1	0.40	94.00
97	1	0.40	94.40
109	1	0.40	94.80
115	1	0.40	95.20
124	1	0.40	95.60
161	1	0.40	96.00

194	1	0.40	96.40
279	1	0.40	96.80
360	1	0.40	97.20
385	1	0.40	97.60
492	1	0.40	98.00
543	1	0.40	98.40
546	1	0.40	98.80
569	1	0.40	99.20
722	1	0.40	99.60
753	1	0.40	100.00
Total	250	100.00	

Figure 12: Frequency, percent and cumulative percent of founded innovative startups.

number_firms

This is the variable indicating the total number of enterprises in a certain province and year and is the denominator in the ratio to compute the dependent variable for the first two analysed models.

Variable	Obs	Mean	Std. Dev.	Min	Max
number_firms	250	51417.53	65012.01	10906	344737

Figure 13: Summary of number of firms' variable.

The minimum and maximum values of this variable, respectively 10906 and 344737, are considerably higher than those observed for startup_prevalenza_femminile and tot_innovative_startup. This difference is due to the fact that this variable includes all types of firms legally registered under Italian regulations between 2013 and 2022, not just female-led or innovative startups. Consequently, the broader scope results in a wider range of values, reflecting the full diversity of firms across provinces and years. The mean of 51417 firms and the high standard deviation of 65012 reflect significant variability in the distribution of firms across different provinces, as for the above variables.

startup_foundation_rate

The startup foundation rate is the dependent variable of the first two models, and it obtained according to the definition given in the Eurostat-OECD Manual on Business Demography Statistics. Therefore, it is calculated as $\frac{startup\_prevalenza\_femminile}{number\_firms}$ and, given the larger dimension of the total number of firms with respect to the innovative startups led by women, the startup_foundation_rate variable assumes small values.

Variable	Obs	Mean	Std. Dev.	Min	Max
startup_fo~e	250	.000093	.0001075	0	.0007328

Figure 14: Summary of female-led startup foundation rate variable.

The summary in Figure 14 shows a minimum of 0 meaning that there is at least one province where no innovative startups led by women were founded in a certain year, in accordance with the minimum value of startup_prevalenza_femminile. The mean value is 0.000093, indicating that, on average, only a very small fraction of firms or entities are newly founded startups in the observed data. Although a standard deviation of 0.0001075 shows some variability in the rate, the values are generally very small.

provincia	Freq.	Percent	Cum.
Alessandria	1	1.05	1.05
Asti	2	2.11	3.16
Bergamo	8	8.42	11.58
Biella	1	1.05	12.63
Brescia	6	6.32	18.95
Como	3	3.16	22.11
Cremona	3	3.16	25.26
Cuneo	5	5.26	30.53
Genova	7	7.37	37.89
Imperia	1	1.05	38.95
La Spezia	2	2.11	41.05
Lecco	2	2.11	43.16
Lodi	5	5.26	48.42
Mantova	3	3.16	51.58
Milano	10	10.53	62.11
Monza e della Brianza	4	4.21	66.32
Novara	7	7.37	73.68
Pavia	8	8.42	82.11
Torino	9	9.47	91.58
Valle d'Aosta/Vallée d'Aoste	3	3.16	94.74
Varese	5	5.26	100.00
Total	95	100.00	

Figure 15: Frequency of provinces with a startup_foundation_rate above the average.

The above table of frequencies shows the frequency of the provinces that satisfy the condition of having a startup_foundation_rate larger than 0.0001, thus greater than the average. Consistent with the findings from the literature revie, Milano is the province attracting the most female entrepreneurs launching innovative startups, with a frequency of 10 indicating that it always has a foundation rate for female innovative startups above the mean. It is followed by Torino, Pavia and Bergamo where female entrepreneurs are most likely to take advantage from research institutions, universities and innovative hubs.
On the contrary, Figure 16 shows the provinces with a rate below the mean. Aligned with what discussed in the literature review, Vercelli, Verbano-Cusio-Ossola, Sondrio and Savona are the provinces with foundation rates below the average during all the observed years.

provincia	Freq.	Percent	Cum.
Alessandria	8	5.37	5.37
Asti	8	5.37	10.74
Bergamo	2	1.34	12.08
Biella	9	6.04	18.12
Brescia	4	2.68	20.81
Como	7	4.70	25.50
Cremona	7	4.70	30.20
Cuneo	5	3.36	33.56
Genova	3	2.01	35.57
Imperia	9	6.04	41.61
La Spezia	8	5.37	46.98
Lecco	8	5.37	52.35
Lodi	5	3.36	55.70
Mantova	5	3.36	59.06
Monza e della Brianza	5	3.36	62.42
Novara	3	2.01	64.43
Pavia	2	1.34	65.77
Savona	10	6.71	72.48
Sondrio	10	6.71	79.19
Torino	1	0.67	79.87
Valle d'Aosta/Vallée d'Aoste	5	3.36	83.22
Varese	5	3.36	86.58
Verbano-Cusio-Ossola	10	6.71	93.29
Vercelli	10	6.71	100.00
Total	149	100.00	

Figure 16: Frequency of provinces with a startup_foundation_rate below the average.

startup_share

The startup foundation share is the dependent variable of the third model, and it calculated as $\frac{startup\_prevalenza\_femminile}{tot\_startup\_innovative}$. It is an indication of the percentage of female-led innovative startups with respect to the total innovative startups.

Variable	Obs	Mean	Std. Dev.	Min	Max
startup_sh~e	250	.2251283	.1894961	0	1

Figure 17: Summary of innovative startup share's variable.

The mean of 0.2251 suggests that approximately 22.5% of the innovative startups in the Nort-West of Italy are led by women, even though this percentage is higher than the one reported by Ministero delle Imprese e del Made in Italy – around 14% - for the general Italian situations. However, the extent of variation from the mean represented by a standard deviation of 0.1895, or 18,95%, indicates that typically female innovative startup share is close to the average. Although the variable also shows values close to the maximum that, in this case, is equal to 1 – 100% of innovative startups are led by women in certain provinces.

Even though it is anticipated that provinces such as Milano and Torino—renowned for their innovative entrepreneurial ecosystems—would show higher shares of female-led startups, the data presented in Figure 18 does not support this expectation. Instead, it reveals that the provinces with a female foundation share exceeding 0.5 are typically those where relatively few innovative startups are founded each year, and these startups happen to be led by women.

provincia	Freq.	Percent	Cum.
Alessandria	1	7.69	7.69
Asti	1	7.69	15.38
Biella	1	7.69	23.08
Imperia	2	15.38	38.46
La Spezia	1	7.69	46.15
Lodi	1	7.69	53.85
Novara	2	15.38	69.23
Pavia	2	15.38	84.62
Valle d'Aosta/Vallée d'Aoste	2	15.38	100.00
Total	13	100.00	

Figure 18: Frequency of provinces with a startup_share lower than 0.5.

When, indeed, summarising the startup foundation shares of cities with the largest foundation rates for women, the figure below shows that it does not necessarily mean that they also have the largest shares of female-led innovative startups. Instead, they exhibit shares close to the average.

startup_share	Freq.	Percent	Cum.
.1	1	3.33	3.33
.1929825	1	3.33	6.67
.2096774	1	3.33	10.00
.2105263	1	3.33	13.33
.2222222	1	3.33	16.67
.2268041	1	3.33	20.00
.2272727	1	3.33	23.33
.2352941	1	3.33	26.67
.2363636	1	3.33	30.00
.2401434	1	3.33	33.33
.2428571	1	3.33	36.67
.25	1	3.33	40.00
.2560976	1	3.33	43.33
.2608696	1	3.33	46.67
.2682927	1	3.33	50.00
.2695652	1	3.33	53.33
.2779221	1	3.33	56.67
.2783505	1	3.33	60.00
.28125	1	3.33	63.33
.3	1	3.33	66.67
.3022847	1	3.33	70.00
.3027523	1	3.33	73.33
.3030303	1	3.33	76.67
.3047091	1	3.33	80.00
.3112339	1	3.33	83.33
.3213812	1	3.33	86.67
.3241758	1	3.33	90.00
.328125	1	3.33	93.33
.3295455	1	3.33	96.67
.3846154	1	3.33	100.00
Total	30	100.00	

Figure 19:Frequency of startup_share of Milano, Torino and Bergamo from 2013 to 2022.

tot_popolazione

The tot_popolazione variable, representing the total population of the provinces between 2013 and 2022 without filters on the gender, nationality and age, is used to compute the rates of the percentage of female foreigners, graduated residents and of the availability of childcare services.

Variable	Obs	Mean	Std. Dev.	Min	Max
tot_popola~e	250	639609	702687.3	123360	3265327

Figure 20: Summary of total populations' variable.

The average population across provinces is 639609. However, a high standard deviation of 702687 shows the considerable variability in the population sizes between provinces, as it is expected in the case of female foreign population. The smallest population observed is 123360, representing the province with the lowest population, while the most populous province has an observed population of 3265327.

donne_immigrate

This variable represents the number of foreign women residents in each province during the years going from 2013 and 2022. It is one of the independent variables of the linear regression models and is employed to analyse the effect of female immigrants on the foundation rate of female immigrants.

Variable	Obs	Mean	Std. Dev.	Min	Max
donne_immi~e	250	25331.36	34051.14	3326	182271

Figure 21: Summary of foreign women's variable.

The summarising table above shows that the variable has a wide interval, going from 3326 to a maximum of 182271 population of female immigrants in a certain province. The mean number of immigrant women per observation is of 25331, while the standard deviation of 34051 indicates a significant variation in the number of immigrant women across observations. The wide range and variability of this regressor can be attributed to its correlation with the diverse total population sizes and geographical areas of the provinces. Provinces with larger land areas in square kilometres tend to have higher total populations, increasing the likelihood of having more immigrant individuals, including foreign women.

The statistics in Figure 21 highlight the four provinces – Bergamo, Milano, Pavia and Torino -with the highest foundation rate for female-led innovative startups, and the ones with the lowest rates – Savona, Sondrio, Verbano-Cusio-Ossola and Vercelli. The respective rate of female immigrants in each province is calculated as $\frac{donne\_immigrate}{tot\_popolazione}$. Considering that this research investigates the impact of increasing the proportion of female immigrants on the foundation rate of female entrepreneurs, the analysis confirms that the rate of female immigrants, rate_donne_immigrate, is higher in the provinces with the most productive startup ecosystems and lower in those with poorer entrepreneurial activities. Therefore, the results from the linear regression models are expected to show a positive coefficient for the regressor representing the proportion of female immigrants, indicating a positive relationship between the two variables.

rate_donne _immigrate	provincia Bergamo	Milano	Pavia	Savona	Sondrio	Torino	Verbano..	Vercelli	Total
.0212088	0	0	0	0	1	0	0	0	1
.0214216	0	0	0	0	1	0	0	0	1
.0220081	0	0	0	0	1	0	0	0	1
.0260511	0	0	0	0	0	0	1	0	1
.0261643	0	0	0	0	0	0	1	0	1
.0265008	0	0	0	0	0	0	1	0	1
.0301553	0	0	0	0	0	0	0	1	1
.0302769	0	0	0	0	0	0	0	1	1
.0307533	0	0	0	0	0	0	0	1	1
.0311144	0	0	0	1	0	0	0	0	1
.0313392	0	0	0	1	0	0	0	0	1
.0324113	0	0	0	1	0	0	0	0	1
.0373227	0	0	0	0	0	1	0	0	1
.0373273	0	0	0	0	0	1	0	0	1
.0373813	1	0	0	0	0	0	0	0	1
.0373906	0	0	0	0	0	1	0	0	1
.0377358	1	0	0	0	0	0	0	0	1
.0384983	1	0	0	0	0	0	0	0	1
.0414811	0	0	1	0	0	0	0	0	1
.0420096	0	0	1	0	0	0	0	0	1
.0424878	0	0	1	0	0	0	0	0	1
.0517895	0	1	0	0	0	0	0	0	1
.0526048	0	1	0	0	0	0	0	0	1
.056225	0	1	0	0	0	0	0	0	1
Total	3	3	3	3	3	3	3	3	24

Figure 22: Summary of rate_donne_immigrate from 2019 to 2021 in Bergamo, Milano, Pavia, Savona, Sondrio, Torino, Verbano-Cusio-Ossola and Vercelli.

tot_servizi_attivi

This is the second independent variable under analysis, and it indicates the total number of facilities for early childhood care services in each province between 2013 and 2022.

Variable	Obs	Mean	Std. Dev.	Min	Max
tot_serviz~i	250	135.064	111.607	26	572

Figure 23: Summary of caregiving services for early childhood's variable.

The smallest number of services observed is 26, while the largest one is 572. The mean of caregiving services per province is 135.06 that, with a standard deviation of 111.61, indicates a notable variation in the number of facilities locally provided across provinces.

rate_servi zi	provincia Bergamo	Milano	Pavia	Savona	Sondrio	Torino	Verbano..	Vercelli	Total
.0000787	0	1	0	0	0	0	0	0	1
.0000839	0	1	0	0	0	0	0	0	1
.0000926	0	1	0	0	0	0	0	0	1
.0001717	0	0	0	0	0	0	1	0	1
.0001855	0	0	0	0	0	0	1	0	1
.0001872	0	0	0	0	0	0	1	0	1
.0002139	1	0	0	0	0	0	0	0	1
.0002256	1	0	0	0	0	0	0	0	1
.0002357	1	0	0	0	0	0	0	0	1
.0002384	0	0	0	0	0	1	0	0	1
.0002412	0	0	0	0	0	1	0	0	1
.0002421	0	0	0	0	0	1	0	0	1
.0002698	0	0	0	0	0	0	0	1	1
.0002775	0	0	0	0	0	0	0	1	1
.0002821	0	0	0	0	0	0	0	1	1
.000432	0	0	1	0	0	0	0	0	1
.000433	0	0	1	0	0	0	0	0	1
.0004386	0	0	1	0	0	0	0	0	1
.0007047	0	0	0	0	1	0	0	0	1
.0007205	0	0	0	0	1	0	0	0	1
.0007361	0	0	0	0	1	0	0	0	1
.0008563	0	0	0	1	0	0	0	0	1
.0008877	0	0	0	1	0	0	0	0	1
.000905	0	0	0	1	0	0	0	0	1
Total	3	3	3	3	3	3	3	3	24

Figure 24: Summary of rate_servizi from 2019 to 2021 in Bergamo, Milano, Pavia, Savona, Sondrio, Torino, Verbano-Cusio-Ossola and Vercelli.

As for the other independent variable, it is provided a more detailed and anticipating analysis on the distribution of the rates of services with respect to the total population and the provinces with the largest and smallest startup foundation rates. Nevertheless, in this case, the analysis does not confirm that an increase in the availability of caregiving services supports an increase in the number of innovative startups founded in a province. Furthermore, the variable does not exhibit a trend in the other direction either, but it reveals an irregular distribution that seems, instead, linked to the dimensions of the area of the provinces. Indeed, larger provinces such as Milano, Torino and Bergamo have the smaller rates in this analysis, suggesting that it is more difficult for local authorities to manage and provide the needed support services when the population and geographic area of the province are large. On the contrary, a smaller square kilometre area, beside having a

smaller number of individuals demanding for caregiving support services, has fewer challenges in allocating the resources for local authorities.
For this reason, it is difficult to predict the effect of the direction of improvement of the independent variable on the regressand and, hence, the sign of the coefficient of the regressor.

residenti_laureati

This is the variable representing the residents that graduated during the years within 2013 and 2022 in the North-Western provinces of Italy. Given that highly skilled and educated individuals are a significant entrepreneurial resource, the number of graduated residents is used as a control variable to account for the impact that the presence of these individuals has on the possibilities to have newly founded innovative startups.

Variable	Obs	Mean	Std. Dev.	Min	Max
residenti_~i	**250**	**2948.716**	**3496.84**	**515**	**18013**

Figure 25: Summary of graduated residents' variable.

As in the case of foreign individuals, also the number of graduated residents depends on the total population and dimension of the area of the province and its subject to their same variability as highlighted by a standard deviation of around 3497 and minimum and maximum value that range from, respectively, 515 and 18013.

Figure 26 provides insight into the distribution of the rate of graduated residents relative to the total population in areas with the highest and lowest foundation rates of female-led innovative startups. Provinces with graduation rates below the average value of 0.0043603 tend to be those with low startup foundation rates. On the contrary, provinces with rates above this mean generally correspond to those with higher foundation rates, although some exceptions are observed for provinces at the lower end of the ranking. Therefore, it is expected that the coefficient for this control variable will show a positive sign in the regression analysis, indicating that increasing the proportion of graduated residents has a positive impact on the foundation rate of innovative startups led by women.

rate_residenti_laureati	provincia Bergamo	Milano	Pavia	Savona	Sondrio	Torino	Verbano..	Vercelli	Total
.0038276	0	0	0	0	0	0	1	0	1
.0039884	0	0	0	0	0	0	0	1	1
.0040046	0	0	0	0	0	0	1	0	1
.0040901	0	0	0	0	1	0	0	0	1
.0041325	0	0	0	0	0	0	0	1	1
.0044642	0	0	0	1	0	0	0	0	1
.004482	0	0	1	0	0	0	0	0	1
.0045054	0	0	0	0	0	0	1	0	1
.0045065	0	0	0	1	0	0	0	0	1
.0045079	0	0	0	0	1	0	0	0	1
.0045394	0	0	1	0	0	0	0	0	1
.0045683	0	0	0	0	0	0	0	1	1
.0045836	0	0	0	0	1	0	0	0	1
.0047451	0	0	0	1	0	0	0	0	1
.0049011	0	0	1	0	0	0	0	0	1
.0049105	0	0	0	0	0	1	0	0	1
.0049667	0	1	0	0	0	0	0	0	1
.0050444	1	0	0	0	0	0	0	0	1
.0051429	1	0	0	0	0	0	0	0	1
.0052365	0	1	0	0	0	0	0	0	1
.0053533	0	0	0	0	0	1	0	0	1
.0055565	0	1	0	0	0	0	0	0	1
.0056403	0	0	0	0	0	1	0	0	1
.0057387	1	0	0	0	0	0	0	0	1
Total	3	3	3	3	3	3	3	3	24

Figure 26: Summary of rate_residenti_laureati from 2019 to 2021 in Bergamo, Milano, Pavia, Savona, Sondrio, Torino, Verbano-Cusio-Ossola and Vercelli.

area_km2

The area in square kilometres of the provinces is used as a control variable to account for all the demographic and economic implications of wider possibilities to launch a startup in a large province.

Variable	Obs	Mean	Std. Dev.	Min	Max
area_km2	250	2317.131	1683.267	405.4018	6894.823

Figure 27: Summary of area in square kilometres' variable.

The average area of the provinces is 2317.13 km^2, with a standard deviation of 1683.27 km^2 indicating a considerable variation in the sizes of the provinces. The smallest observed area is 405.40 km^2, representing the province with the least land area, that is Monza e della Brianza. Whereas the largest observed area is 6894.82 km^2, corresponding to the province of Cuneo.

value_added_per_inhabitant

The value added per inhabitant is used as control variable to account for a measure of economic productivity across the provinces in the observed years.

Variable	Obs	Mean	Std. Dev.	Min	Max
value_adde~t	225	64413.37	5290.258	54347.9	85957.15

Figure 28: Summary of value added per inhabitant's variable.

For this variable, the summarising table only shows 225 observations in total due to the missing values for the year 2022 from the original database. The average per capita value added is 64413.37, with a relatively small standard deviation of 5290.26 suggesting moderate variability in value added per capita across provinces. The lowest observed value added per inhabitant is 54347.90 in Verbano-Cusio-Ossola, while the highest value is 85957.15 in Milano, indicating that it is the most economically productive province.

value_added_per_inhabitant	provincia								
	Bergamo	Milano	Pavia	Savona	Sondrio	Torino	Verbano..	Vercelli	Total
55953.39	0	0	0	0	0	0	1	0	1
60428.14	0	0	0	0	1	0	0	0	1
60492.28	0	0	0	0	0	0	1	0	1
60575.26	0	0	1	0	0	0	0	0	1
61477.06	0	0	0	1	0	0	0	0	1
61493.33	0	0	0	0	0	0	0	1	1
61690.04	0	0	0	0	0	0	1	0	1
61722.44	0	0	0	1	0	0	0	0	1
62043.18	0	0	0	0	0	0	0	1	1
62228.43	0	0	0	0	0	1	0	0	1
63275.68	0	0	0	0	1	0	0	0	1
64055.61	0	0	0	0	1	0	0	0	1
64341.75	0	0	1	0	0	0	0	0	1
65205.49	1	0	0	0	0	0	0	0	1
66216.29	0	0	1	0	0	0	0	0	1
66413.34	0	0	0	0	0	1	0	0	1
66943.93	0	0	0	1	0	0	0	0	1
66993.41	0	0	0	0	0	0	0	1	1
67402.93	0	0	0	0	0	1	0	0	1
68868.97	1	0	0	0	0	0	0	0	1
71432.46	1	0	0	0	0	0	0	0	1
78488.05	0	1	0	0	0	0	0	0	1
80106.84	0	1	0	0	0	0	0	0	1
85957.15	0	1	0	0	0	0	0	0	1
Total	3	3	3	3	3	3	3	3	24

Figure 29: Summary of value_added_per_inhabitant from 2019 to 2021 in Bergamo, Milano, Pavia, Savona, Sondrio, Torino, Verbano-Cusio-Ossola and Vercelli.

As in the previous statistics, the table in Figure 29 provides information about the underlying relationship between the value added per capita and the foundation rate of female-led innovative startups. While the extreme values, both lowest and highest, seem to confirm that higher value added per capita is associated with higher startup foundation rates, the rest of the analysed data does not suggest a specific direction of improvement. Consequently, interpreting the effect of an increase in value added per capita is not straightforward. However, given its economic relevance, there are expectations for a positive coefficient for this variable in the regression model.

unemployment_rate

The unemployment rate is used as a control variable to represent the labour market conditions of the provinces.

Variable	Obs	Mean	Std. Dev.	Min	Max
unemployme~e	250	7.579497	2.18469	2.815521	14.44494

Figure 30: Summary of unemployment rate's variable.

The key statistics reveal a mean unemployment rate of 7.58% with a moderate standard deviation of 2.18%. The lowest observed rate is 2.82%, while the highest reaches 14.44%, indicating significant unemployment in the province with the highest value. Specifically, the lowest rates are found in provinces within the region of Lombardia, while the highest rates are concentrated in regions like Liguria, particularly in Imperia, and Piemonte.

unemployme	provincia								
nt_rate	Bergamo	Milano	Pavia	Savona	Sondrio	Torino	Verbano..	Vercelli	Total
3.07	1	0	0	0	0	0	0	0	1
3.54	1	0	0	0	0	0	0	0	1
3.55	1	0	0	0	0	0	0	0	1
5.44	0	0	0	0	1	0	0	0	1
5.60	0	0	1	0	0	0	0	0	1
5.62	0	0	0	0	1	0	0	0	1
5.68	0	0	0	1	0	0	0	0	1
5.81	0	0	0	0	0	0	1	0	1
5.85	0	0	0	0	0	0	1	0	1
5.92	0	1	0	0	0	0	0	0	1
5.94	0	1	0	0	0	0	0	0	1
6.18	0	0	0	0	0	0	1	0	1
6.42	0	0	0	1	0	0	0	0	1
6.47	0	1	0	0	0	0	0	0	1
6.48	0	0	0	0	1	0	0	0	1
6.62	0	0	1	0	0	0	0	0	1
6.95	0	0	1	0	0	0	0	0	1
7.57	0	0	0	1	0	0	0	0	1
7.87	0	0	0	0	0	0	0	1	1
8.23	0	0	0	0	0	0	0	1	1
8.23	0	0	0	0	0	0	0	1	1
8.26	0	0	0	0	0	1	0	0	1
8.33	0	0	0	0	0	1	0	0	1
8.33	0	0	0	0	0	1	0	0	1
Total	3	3	3	3	3	3	3	3	24

Figure 31: Summary of unemployment_rate from 2019 to 2021 in Bergamo, Milano, Pavia, Savona, Sondrio, Torino, Verbano-Cusio-Ossola and Vercelli.

The statistics in Figure 31 summarise the unemployment rate for the provinces with the highest and lowest values of foundation rate. Although it is expected to see higher unemployment rates in provinces with issues in the labour market conditions and, thus, with less opportunities for female entrepreneurs to launch an innovative startup, the data on the table show a diverse scenario. Indeed, lower unemployment rates are not necessarily connected to the provinces with the highest foundation rates and vice versa. Consequently, these statistics do not anticipate the direction of the effect of an increase in the unemployment rate on the foundation rate of female-led innovative startups and, thus, the sign of the control variable's coefficient. However, a high unemployment rate in provinces with a high startup foundation rate might be an indicator of necessity entrepreneurship, suggesting that inappropriate market labour conditions encourage entrepreneurial activities led by women.

5. Results

This chapter reports and describes the results of the three statistical models that test the hypothesis stated in Paragraph 4.2.

5.1 Model 1

The first hypothesis aims to test the effect of an increase in the proportion of foreign women as a labour force in support to local caregiving services on the foundation rate of female-led innovative startups. Based on the model described in Paragraph 4.2, a linear regression analysis was conducted using startup_foundation_rate as the dependent variable. The independent variables, or regressors, include rate_donne_immigrate - proportion of immigrant women- and rate_servizi - availability of caregiving services for early childhood. Additionally, five control variables were included to account for other influential factors: rate_residenti_laureati, that is the rate of residents with a university degree, ln_value_added_inhabitant, the linear-log distribution of value-added per inhabitant, unemployment_rate, area_km2 and tot_popolazione. These control variables enable the isolation of the effect of the independent variables on the startup foundation rate by accounting for demographic, economic, and geographic variations across provinces.

5.1.1 Model Fit and Performance

Source	SS	df	MS			
				Number of obs	=	225
				F(7, 217)	=	48.49
Model	1.6144e-06	7	2.3062e-07	Prob > F	=	0.0000
Residual	1.0321e-06	217	4.7562e-09	R-squared	=	0.6100
				Adj R-squared	=	0.5974
Total	2.6465e-06	224	1.1815e-08	Root MSE	=	6.9e-05

Figure 32: Analysis of the fit of Model 1.

Figure 32 presents an overview of the model's fit, illustrating how well the selected variables and the estimated regression align with the actual data and explain the relationship under analysis.

The analysis of the variance in the table on the left shows the portion of variance explained by each of the sources, thus by the model, the residuals and the sum of these two. From the values of the Sum of Squares SS it is possible to notice that the variance explained by the model is equal to 1.6144e-06, whereas the unexplained variance and so the one left to the residuals is equal to 1.0321e-06. The second column, instead, report the Degrees of Freedom df that, as discussed in the chapter explaining the Empirical Method, are equal to the number of observations minus 1, for a total of 224. The last column of the table shows the Mean Squares MS, that are computed as the SS divided by the df for each source. Although the dataset was designed to include observations from 25 provinces over 10 years - from 2013 to 2022 - amounting to a total of 250 observations, the data presented in the figure above indicates only 225 total observations. This reduction is due to missing data for the year 2022 in the control variable ln_value_added_per_inhabitant. The software used to perform the linear regression analysis automatically excludes any observations with missing values across any of the variables included in the model, ensuring that the results are not distorted by incomplete data. The p-value associated with the F-statistics, that is computed as the ratio of the model MS to the residual MS, suggests that the model is statistically significant given that $p < 0.05$.

The measures of fit considered to evaluate the model are the ones described in Chapter 3 and reported on the right side of the table. Starting from the R-squared, its value suggests that the proportion of variance in the dependent variable explained by the model is equal to a relatively high value equal to 61%, meaning that the independent and control variables are good predictors of the regressand. However, to provide a more accurate assessment of model fit, especially in linear regressions with multiple predictors, the Adjusted R-squared is calculated. In this case, the value of 59.74% confirms that the model maintains a strong fit, even after adjusting for the number of regressors. The last measure of model fit analysed is the Root Mean Squared Error, or SER, and it reports a measure of the model's prediction error. With a value of 6.9e-05 close to 0, it supports the validity of the model fits already exhibited by the R^2 and Adjusted R^2, indicating a low standard deviation of the residuals of the regression.

Variable	VIF	1/VIF
ln_value_a~t	3.32	0.301607
rate_resid~i	2.73	0.366546
tot_popola~e	2.27	0.440196
rate_donne~e	1.87	0.533609
unemployme~e	1.36	0.734007
area_km2	1.36	0.734222
rate_servizi	1.27	0.786168
Mean VIF	2.03	

Figure 33: VIF and Tolerance Values for Independent Variables in the Regression Model.

The table in Figure 33 describes the Variance Inflation Factor, or VIF, values and their reciprocals for the set of independent and control variables used in a regression model. These are measures that help to identify multicollinearity among the variables and, generally, higher VIF values suggest higher multicollinearity. In this case, none of the variables shows a VIF value that exceeds the typical limit of 10 and, with the highest being 3.32, it indicates that multicollinearity is not a significant concern in this model.

5.1.2 Analytical Results

For a comprehension of the effect of the regressors on the dependent variable, this paragraph analyses and interprets the output of the linear regression analysis.

Before analysing the coefficients, it is important to note that all independent variables, except for area in square kilometres, total population, and value added per inhabitant, are expressed as rates relative to the total population of each province. The variable value added per inhabitant is further transformed into its logarithmic form to mitigate any distortions caused by its variability. To correctly understand and explain the impact of each regressor on the startup foundation rate, it is necessary to keep in mind the type of distribution of each of the variables.

startup_foundation_rate	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rate_donne_immigrate	.0032585	.0008983	3.63	0.000	.001488	.0050289
rate_servizi	.0115259	.0255692	0.45	0.653	-.0388699	.0619216
rate_residenti_laureati	.0303588	.0131579	2.31	0.022	.0044251	.0562925
unemployment_rate	-5.04e-06	2.52e-06	-2.00	0.047	-.00001	-7.34e-08
ln_value_added_inhabitant	.0002861	.0001052	2.72	0.007	.0000788	.0004934
area_km2	-3.37e-09	3.19e-09	-1.06	0.293	-9.67e-09	2.92e-09
tot_popolazione	6.25e-11	9.88e-12	6.32	0.000	4.30e-11	8.19e-11
_cons	-.0033183	.0011213	-2.96	0.003	-.0055283	-.0011084

Figure 34: Results of the linear regression of Model 1.

The table above shows the results of the linear regression of Model 1. Given that the proportion of female led-innovative startups is definitely small compared to the total number of firms in a province, the output of the analysis and, specifically, the coefficients of the regressors are difficult to interpret. To facilitate the discussion, the same linear regression was computed with the dependent variable startup_foundation_rate multiplied by 1000. This adjustment quantifies the proportion of innovative startups led by women per 1000 firms in the province. The resulting output, shown below, reflects the same scenario as in Figure 34 but presents the coefficients in a more interpretable and readable format.

startup_foundation_rate	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rate_donne_immigrate	3.25848	.8982673	3.63	0.000	1.488034	5.028925
rate_servizi	11.52586	25.56921	0.45	0.653	-38.86993	61.92165
rate_residenti_laureati	30.35881	13.15791	2.31	0.022	4.425142	56.29247
unemployment_rate	-.0050389	.0025194	-2.00	0.047	-.0100045	-.0000734
ln_value_added_inhabitant	.2861378	.105182	2.72	0.007	.0788287	.4934468
area_km2	-3.37e-06	3.19e-06	-1.06	0.293	-9.67e-06	2.92e-06
tot_popolazione	6.25e-08	9.88e-09	6.32	0.000	4.30e-08	8.19e-08
_cons	-3.318319	1.121253	-2.96	0.003	-5.52826	-1.108378

Figure 35: Results of the linear regression of Model 1 - version 2.

With reference to the first independent variable rate_donne_immigrate, representing the proportion of immigrant women, its positive coefficient highlights a positive relation with the foundation rate of innovative startups led by women. In particular, a 1-unit increase in the proportion of immigrant women is associated with 3.26 more female-led startups per 1000 firms. The related p-value of the regressor verifies the null hypothesis testing that the two variables are not statistically significant. Given a p-value lower than 0.05, it is possible to reject the null hypothesis and to confirm that the positive influence of the ratio of female immigrants on the startup foundation rate is statistically significant. The other independent variable, representing the ratio of local services for early childhood, shows a positive coefficient suggesting that a 1-unit increase in the proportion of services would result in 11.53 more female-led innovative startups per 1000 firms. In contrast with the previous case, this result is statistically insignificant as its p-value is equal to 0.653, so the relationship cannot be considered meaningful. As expected from the descriptive statistics analysis in Chapter 4, the proportion of graduated residents has a positive relation with female-led innovative businesses. Precisely, a 1-unit increase in the proportion of residents with a university degree is associated with 30.36 more female-led startups per 1000 firms. Its significance is verified by a p-value of 0.022, validating the findings from the literature review saying that, for female entrepreneurship, the level of education is relevant. The unemployment rate has a small but significant negative effect on the dependent variable, suggesting that innovative female-led startups in the North-Western provinces of Italy are not raising from necessity entrepreneurship, but from business opportunities. A 1% increase in unemployment is associated with a decrease of about 0.005 female-led startups per 1000 firms. The log-transformation of value added per inhabitant is significantly and positively related to the startup foundation rate, meaning that higher economic output per inhabitant encourages the formation of female-led startups. Precisely, a 1% increase in value added per inhabitant is associated with an increase of 0.286 female-led startups per 1000 firms. Whereas, the total population has a positive and significant effect, indicating that for every additional inhabitant, approximately 6.25e-08 more female-led startups per 1000 firms are founded, the area has a negative but statistically insignificant impact, meaning that the size of the province does not play a major role in the startup foundation rate. The constant term is significant but negative and, therefore, it only has a mathematical interpretation.

5.2 Model 2

The second model tests the second hypothesis stated in Chapter 4. In addition to evaluating the impact of the independent and control variables already presented in Model 1, this model includes an interaction term aiming to assess whether the combined effect of an increase in the proportion of female foreigners and of caregiving services within a province has a significant influence on the foundation rate of female-led startups.

To simplify to readability and interpretation of results, the dependent variable considered for the analysis is $\frac{startup\_prevalenza\_femminile}{number\_firms} * 1000$, therefore, referring to the foundation rate of innovative startups led by women per 1000 firms.

5.2.1 Model Fit and Performance

Source	SS	df	MS			
				Number of obs	=	225
				F(8, 216)	=	43.76
Model	1.63665472	8	.204581841	Prob > F	=	0.0000
Residual	1.009802	216	.004675009	R-squared	=	0.6184
				Adj R-squared	=	0.6043
Total	2.64645672	224	.011814539	Root MSE	=	.06837

Figure 36: Analysis of the fit of Model 2.

As with the previous model, the results of the linear regression analysis are based on a total of 225 observations, as shown in the table above. The variance explained by the model and the residual variance are similar to those in Model 1, while the degrees of freedom show a slight difference. Although the total degrees of freedom remain at 224, the model uses one additional degree of freedom in Model 2 to evaluate the interaction term included as an additional regressor. The F-statistic reported on the right measures the overall significance of the model, and its high value of 43.76 together with the p-value = 0.000 indicate that the model is statistically significant. With regards to the three fit measures of the model, the R-squared and Adjusted R-squared only report a minor improvement, with respectively 61.84% and 60.43% of the dependent variable being explained by the predictors used in the model. The same reasoning is applied to the Root Mean Squared Error that measures the average distance between the observed and predicted values. Its low value confirms the good fit of the model.

Variable	VIF	1/VIF
rate_donne~e	6.10	0.163809
rate_servizi	26.15	0.038241
rate_resid~i	2.91	0.343218
unemployme~e	1.36	0.732819
ln_value_a~t	3.58	0.279018
area_km2	1.37	0.728425
tot_popola~e	2.58	0.387314
c. rate_donne~e# c. rate_servizi	23.66	0.042263
Mean VIF	8.47	

Figure 37: VIF and Tolerance Values for Independent Variables in the Regression Model.

This table presents the Variance Inflation Factor and its inverse for the independent variables in the regression model, which includes an interaction term between rate_donne_immigrate and rate_servizi. The first variable has a moderate VIF, suggesting some multicollinearity, but generally acceptable since it's below 10. The high VIF values for rate_servizi and its interaction term with rate_donne_immigrate suggest significant multicollinearity, which may distort the precision of the coefficient estimates for these variables. The other variables, instead, show acceptable levels of multicollinearity. Nevertheless, the average VIF of the model is 8.47, a relatively high number even though it is below 10. Consequently, it might be necessary to understand how to address the multicollinearity of rate_servizi and of the interaction term to obtain a non-distorted model.

5.2.2 Analytical Results

The table below presents the results of a linear regression analysis, including the interaction term between rate_donne_immigrate and rate_servizi, which tests whether the effect of immigrant women on female-led startup foundation varies depending on the level of caregiving services in a province.

startup_foundation_rate	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rate_donne_immigrate	6.180352	1.607351	3.85	0.000	3.012252	9.348452
rate_servizi	256.3307	114.9396	2.23	0.027	29.78382	482.8776
rate_residenti_laureati	37.78531	13.48118	2.80	0.006	11.2138	64.35681
unemployment_rate	-.0048194	.0024998	-1.93	0.055	-.0097465	.0001078
ln_value_added_inhabitant	.2213473	.1084195	2.04	0.042	.0076517	.4350429
area_km2	-3.99e-06	3.18e-06	-1.25	0.211	-.0000103	2.28e-06
tot_popolazione	5.46e-08	1.04e-08	5.23	0.000	3.40e-08	7.51e-08
c.rate_donne_immigrate#c.rate_servizi	-7898.737	3617.256	-2.18	0.030	-15028.38	-769.0977
_cons	-2.726025	1.144257	-2.38	0.018	-4.981363	-.4706869

Figure 38: Results of the linear regression of Model 2.

Starting from the independent variables rate_donne_immigrate and rate_servizi, they both have a positive coefficient suggesting an individually positive relation with the foundation rate of innovative startups. Indeed, a unit increase in the proportion of immigrant women in the population is associated with an increase of approximately 6.18 female-led startups per 1000 firms. Whereas a unit increase in the proportion of caregiving services for early childhood relative to the population is associated with a significant increase of approximately 256.33 female-led startups per 1000 firms. Both coefficients are statistically significant with a p-value lower than 0.05, although the second one is rather large and it could be due to the interaction with rate_donne_immigrate, as highlighted by its VIF. In accordance with the model that does not include the term accounting for the combined effect of the independent variables, the control variables rate_residenti_laureati, ln_value_added_inhabitant and tot_popolazione report positive and statistically significant coefficients. In particular, a unit increase in the proportion of residents with university degrees is associated with an increase of 37.79 female-led startups per 1000 firms, a 1% increase in value added per inhabitant is associated with an increase of 0.221 startups per 1000 firms and an additional person in the province increases the female-led startup rate by 5.46e-08 startups per 1000 firms. In contrast, the area in square kilometres of the province has a negative but insignificant effect on the regressand. While the unemployment_rate variable is marginally significant with a p-value = 0.055 and, so, a 1% increase in the unemployment rate is associated with a slight reduction in the female-led startup rate by approximately 0.0048 startups per 1000 firms.

The negative interaction term suggests that while both rate_donne_immigrate and rate_servizi have positive individual effects, their combined effect is negative. This means that in provinces where both the proportion of immigrant women and caregiving services rises, the expected increase in female-led startups is smaller than when these factors are considered individually and, in fact, it seems to decrease the possibilities to launch an innovative startup. This interaction is statistically significant, indicating a meaningful relationship. However, as for rate_servizi, the coefficient is a rather large value that needs further investigation, especially knowing about its multicollinearity with the two independent variables. The constant is negative and significant, even though it does not have a meaningful economic interpretation.

5.3 Model 3

The third analysed model tests the impact of female immigrants and of caregiving services on the share of the innovative startups founded by women with respect to the ones founded by men. The independent and control variables used in this analysis are consistent with those in Model 1. However, the main change lies in the definition of the dependent variable, which is now calculated as $\frac{startup\_prevalenza\_femminile}{tot\_startup\_innovative}$. By altering the calculation method, the dependent variable now reflects the dimension of the share rather than its size relative to the overall businesses. In this way the model provides additional insights on the representation of female entrepreneurs in innovative startups.

Given that the ratio in the dependent variable allows for greater readability compared to the previous case, it is not necessary to multiply by 1000 to interpret the results.

5.3.1 Model Fit and Performance

Source	SS	df	MS			
				Number of obs	=	225
				F(7, 217)	=	4.66
Model	.953624916	7	.136232131	Prob > F	=	0.0001
Residual	6.33928433	217	.029213292	R-squared	=	0.1308
				Adj R-squared	=	0.1027
Total	7.29290924	224	.032557631	Root MSE	=	.17092

Figure 39: Analysis of the fit of Model 3.

The results presented in the table in Figure 39 summarise the fit of the regression analysis with 225 observations of Model 3. The R-squared value indicates that approximately 13.08% of the variability in the dependent variable can be explained by the independent variables in the model. This suggests a weak explanatory function of the regressors, implying that additional or different variables may be needed to better capture the underlying relationships. The adjusted R-squared is slightly lower at 10.27%, reinforcing the possibility that the model may not be optimally explained and that, indeed, these indicators penalise the inclusion of unsuitable variables. Additionally, the RSE is reported at 0.17092, suggesting that the model's predictions are not as precise as in Model 1 and Model 2.

In general, the low values of model fit can be attributed to the formula employed to calculate the startup foundation rate of female-led innovative businesses. This model captures a dimension that focuses on the increase in the share of female-led startups rather than their size within the overall landscape of innovative businesses. Consequently, this emphasis on relative proportions rather than absolute sizes may limit the model's ability to fully explain the variability in the dependent variable, thus resulting in lower R-squared and adjusted R-squared values. This suggests that while the model identifies trends in representation, it may not effectively account for the critical factors influencing the decision and opportunity to launch a business.

Variable	VIF	1/VIF
ln_value_a~t	3.32	0.301607
rate_resid~i	2.73	0.366546
tot_popola~e	2.27	0.440196
rate_donne~e	1.87	0.533609
unemployme~e	1.36	0.734007
area_km2	1.36	0.734222
rate_servizi	1.27	0.786168
Mean VIF	2.03	

Figure 40: VIF and Tolerance Values for Independent Variables in the Regression Model.

The table reporting the Variance Inflation Factor and their reciprocals for the set of independent and control variables used in a regression model, shows that same results as for Model 1 since the same independent and control variables are used for the analysis. Therefore, none of the variables has a VIF higher than 10, or even 6, and, apparently, there is no need to account for multicollinearity in the model.

5.3.2 Analytical Results

startup_foundation_rate	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rate_donne_immigrate	8.264674	2.22621	3.71	0.000	3.876911	12.65244
rate_servizi	-6.423442	63.36915	-0.10	0.919	-131.3213	118.4744
rate_residenti_laureati	83.66465	32.60975	2.57	0.011	19.39226	147.9371
unemployment_rate	.0069022	.0062439	1.11	0.270	-.0054042	.0192086
ln_value_added_inhabitant	.0963987	.2606765	0.37	0.712	-.4173833	.6101807
area_km2	3.74e-06	7.92e-06	0.47	0.637	-.0000119	.0000193
tot_popolazione	-3.57e-08	2.45e-08	-1.46	0.146	-8.40e-08	1.25e-08
_cons	-1.54559	2.778845	-0.56	0.579	-7.022572	3.931392

Figure 41: Results of the linear regression of Model 3.

In accordance with the description of the variability of the model in Paragraph 5.3.1, it is immediately apparent from the p-values that the regressors do not explain this definition of startup foundation rate that refers more to the share rather than the size of innovative businesses led by women.

Among the independent variables, the positive coefficient for rate_donne_immigrate indicates that a unit increase in the proportion of foreign women over the total population is associated with an increase of 8.26 female-led startups per 1000 firms, with this being one of the only two statistically significant variables. Conversely, rate_servizi has a negative coefficient of -6.42, with a statistically non-significant p-value that suggests no meaningful impact on the startup foundation rate. rate_residenti_laureati is the only other regressor being statistically significant, and it is positively associated with the dependent variable in a way that a unit increase in the proportion of graduated residents is associated with an increase of around 84 female-led businesses. The other control variables, such as unemployment_rate, ln_value_added_inhabitant, area_km and tot_popolazione show no significant relationships, with p-values of 0.270, 0.712, 0.637 and 0.146 respectively. The constant term of -1.55, while not having a meaningful economic explanation, it is also not statistically significant.

In general, these findings validate the fact that the definition of startup foundation rate, allowing for an analysis of the factors that influence it, is the one where the number of innovative startups led by women is divided by the total number of firms. Additionally, the results suggest the involvement of other variables to investigate the factors that promote female entrepreneurship with respect to only innovative startups.

6. Discussion and Limitations

The purpose of this thesis is to investigate the role of foreign women and of local caregiving services on the entrepreneurial activity of women in terms of foundation rate of innovative female-led startups across the provinces of the North-West of Italy. To achieve this objective, three linear regression models were constructed and analysed observations from 2013 to 2022 to explore and understand the underlying relationships between the factors influencing the foundation rate of female-led innovative startups.

The three statistical models provide valuable insights, although with some differences in the significance and magnitude of the results. In all three models the variable representing the proportion of immigrant women consistently shows a positive and statistically significant relationship with female-led startups foundation rate. This highlights the important role immigrant women play in driving female entrepreneurship. In contrast, the variable capturing the proportion of caregiving services, shows mixed results. While its coefficient is positive in two models and even significant in the second one, its statistically non-significant value in the other models raises concerns about the potential multicollinearity and the need for further research into its real influence.

The interaction between the rate of female immigrants and the rate of services in the model incorporating the interaction term reveals a negative combined effect, indicating that the simultaneous rise in immigrant women and caregiving services may obstacle the expected increase in female-led startups. This unexpected finding suggests the possibility of an underlying competition between these two factors in the analysed geographic area. It could imply that while both immigrant women and caregiving services independently support female entrepreneurship, when both factors increase together, the resources or opportunities available might not be effectively distributed to promote female-led startup activity. However, further research is needed to better comprehend the dynamics causing a high multicollinearity among these variables. Among the control variables, the proportion of residents with university degrees, consistently shows a positive and significant impact on female-led startup formation across all models, reinforcing the importance of education in promoting female entrepreneurship. Other control variables, such as the value added per inhabitant and the total population, generally exhibit positive and significant relationships with the dependent variable, indicating that economic productivity and population size

support the formation of innovative female-led startups. In contrast, the unemployment rate shows a small and negative effect in all models, which is either marginally significant or insignificant, suggesting that female-led innovative startups in these regions may not arise from necessity-driven entrepreneurship but are rather sustained by opportunity. The geographic size of the province, instead, remains consistently negative and insignificant, indicating that the spatial dimension of a province does not have a major influence.
The last model, which focuses on the startup foundation rate as a share rather than size, seems to not accurately predict the dependent variable, with several regressors becoming insignificant. This outcome suggests that the factors influencing female entrepreneurship may differ when looking at relative representation rather than total entrepreneurial activity. Therefore, the first two models are more reliable and, especially, accurate in the explanation of the researched context.

6.1 Limitations

Despite the valuable insights gained from the analysis, this study has several limitations that must be acknowledged. First, the geographic focus on the North-West of Italy presents a significant constraint. This area includes two provinces, Milano and Torino, with particularly strong and dynamic entrepreneurial ecosystems, which may disproportionately influence the results. Consequently, the findings cannot be generalized to represent the startup foundation rate for female-led innovative businesses across Italy, given the highly fragmented nature of the entrepreneurial landscape across the different Italian regions. Furthermore, the time frame of the study, from 2013 to 2022, includes two critical events: the introduction of the Italian Startup Act and the economic changes caused by the Covid-19 pandemic in 2020. Both events significantly altered the startup environment, particularly by creating more opportunities for innovative businesses. Therefore, the effects of these events may have impacted the outcomes of the linear regression models, potentially altering the results.
Another limitation concerns the restricted focus on caregiving services for early childhood due to the unavailability of data for other forms of caregiving, such as elder care, at the provincial level. This data limitation may account for the anomalous and possibly misleading coefficient values for the variable representing caregiving services, as the analysis lacks comprehensive information of the full panorama of caregiving resources available.

Financial conditions are critical to evaluate opportunities to launch an innovative startup, but these aspects remain unexplored in this study. Indeed, the absence of variables related to the financial situation of the startups, such as debt incurred at the time of founding or access to financial resources, limits the scope of the analysis.

Finally, the lack of data on key characteristics of the entrepreneurs, such as age, nationality, and education level, represents a further limitation. The literature suggests that these factors, particularly age, which often correlates with experience, can significantly influence the decision to launch an innovative startup. Therefore, the exclusion of such variables may ignore important determinants of entrepreneurial behaviour that could help refine and improve the accuracy of the models used in this research.

7. Conclusions

Based on the findings of this research, it is evident that female immigrants have a positive impact on the foundation rate for female-led innovative startups in the North-West of Italy. This suggests that immigrant women contribute to alleviating domestic responsibilities for native female entrepreneurs, allowing them to engage in entrepreneurial activities. However, while the presence of caregiving services seems to have a positive but non-significant effect, the interaction between immigrant women and caregiving services shows a complex dynamic, revealing that the simultaneous increase in both factors can have a negative combined effect, potentially due to competition for resources. The study also demonstrates the importance of education and economic conditions, as higher education levels and economic productivity positively impact the foundation of female-led startups. In contrast, the analysis suggests that unemployment does not significantly drive necessity-based entrepreneurship in these regions, which signals that female-led innovative startups are predominantly opportunity-driven.

Future research could benefit from addressing the limitations of this study. First, including data on other forms of caregiving facilities, such as elder care, would provide a more comprehensive understanding of the impact of caregiving services on female entrepreneurship and, potentially, a statistically significant factor. Additionally, expanding the geographic focus beyond the North-West of Italy would allow for more generalised findings applicable to the fragmented entrepreneurial situation of the country. Finally, the inclusion of financial conditions, such as startup capital or access to funding, and of more detailed demographic data about the entrepreneurs, such as age and nationality, would also enhance the robustness of the analysis and refine the model.

References

- Abouzahr K., Krentz M., Harthorne J. & Brooks Taplett F. (2018). Why Women-Owned Startups Are a Better Bet. BCG (Boston Consulting Group).
- Acs, Z.J., Audretsch, D.B. & Lehmann, E.E. (2013) The knowledge spillover theory of entrepreneurship. Small Bus Econ 41, 757–774.
- Acs, Z.J., Audretsch, D.B. & Lehmann, E.E. (2013). The knowledge spillover theory of entrepreneurship. Small Bus Econ 41, 757–774.
- Amorós J., Cristi O., Naudé W. (2021). Entrepreneurship and subjective well-being: Does the motivation to start-up a firm matter? Journal of Business Research, 127, 389-398.
- Amorós, J. E., Ciravegna, L., Mandakovic, V., Stenholm, P. (2019). Necessity or Opportunity? The Effects of State Fragility and Economic Development on Entrepreneurial Efforts. Entrepreneurship Theory and Practice, 43, 725-750.
- Arcuri M.C., Gandolfi G., Russo I. (2023). Exploring the impact of innovation intensity on the growth of female-led entrepreneurial firms. Journal of Small Business and Enterprise Development.
- Audretsch D., Belitski M., Brush C. (2022). Innovation in women-led firms: an empirical analysis. Innovative behavior of minorities, women and immigrants, 1-21.
- Audretsch D., Colombelli A., Grilli L., Minola T., Rasmussen E. (2020). Innovative start-ups and policy initiatives, Research Policy 49.
- Autio E. & Heikki R. (2016). Retaining Winners: Can Policy Boost High-Growth Entrepreneurship? Research Policy 45, 42-55.
- Barone G., Mocetti S. (2011). With a Little Help from Abroad: The Effect of Low-skilled Immigration on Female Labour Supply. Labour Economics.
- BCG (Boston Consulting Group) & SISTA. (2023). Women-led startups losing across the board: from creation to funding, in all key European markets. BCG (Boston Consulting Group).

- Brunello G., Lodigiani E., Rocco L. (2020). Does low-skilled immigration increase profits? Evidence from Italian local labour markets. Regional Science and Urban Economics 85.
- Cantamessa M., Montagna F. (2016). Management of Innovation and Product Development: Integrating Business and Technology Perspectives, 5-7.
- Colombelli A., D'Ambrosio A., Ravetti C. (2024). Women in innovative start-ups and regional inclusiveness: 'green' and socially-responsible companies. Regional Studies. 1-14.
- Colombelli A. (2016). The impact of local knowledge bases on the creation of innovative start-ups in Italy. Small Bus Econ 47, 383–396.
- Colombelli, A., Krafft, J. & Vivarelli, M. (2016). To be born is not enough: the key role of innovative start-ups. Small Bus Econ 47, 277–291.
- Cortes P., Pan J. (2019). When Time Binds: Substitutes for Household Production, Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled. Journal of Labor Economics.
- Cortes, P. (2008). The Effect of Low-skilled Immigration on US Prices: Evidence from CPI Data. Journal of Political Economy.
- Cortes P., Tessada, J. (2011). Low-skilled immigration and the labor supply of highly skilled women. American Economic Journal: Applied Economics.
- Crunchbase. (2023). 2023 Global Startup Ecosystem: Insights on funding and innovation. Retrieved from https://about.crunchbase.com/blog/2023-global-startup-ecosystem/
- D'Ambrosio, A., Gabriele, R., Schiavone, F. et al. (2017). The role of openness in explaining innovation performance in a regional context. J Technol Transf 42, 389–408. https://doi.org/10.1007/s10961-016-9501-8
- Dealroom.co. (2024). Women-founded startups in Europe. https://dealroom.co/guides/women-founded-startups-europe
- Del Bosco, B., Mazzucchelli, A., Di Gregorio, A., & Chierici, R. (2019). Innovative startup creation: The effect of local factors and demographic characteristics of entrepreneurs. Journal of Small Business & Entrepreneurship, 32, 579-605.

- European Commission. (2014). Commission proposes measures to improve long-term competitiveness of the EU's single market.
- European Commission. (2023). Employment and Social Developments in Europe 2023. Publications Office of the European Union.
- European Commission: Directorate-General for Research and Innovation, Hollanders, H., Es-Sadki, N. and Khalilova, A., European Innovation Scoreboard 2022, Publications Office of the European Union, 2022.
- Farré, L., González, L. & Ortega, F. (2011). Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women. The B.E. Journal of Economic Analysis & Policy, 11(1).
- Forrer V., Santini E., Kuebart A., Kosters T. (2023). Startup Heatmap Europe. Deep Ecosystems
- GEM (Global Entrepreneurship Monitor) (2023). Global Entrepreneurship Monitor 2022/23 Women's Entrepreneurship Report.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. American Economic Review, 104(4), 1091-1119.
- Grilli, L., Mrkajic, B. & Giraudo, E. (2023) Industrial policy, innovative entrepreneurship, and the human capital of founders. Small Bus Econ 60, 707–728.
- Infocamere. (2014). Definizione imprese femminili. Retrieved from https://parsmodir.com/wp-content/uploads/2019/08/InnEcho-Moore1993.pdf
- Istat. (2024). Cittadini non comunitari regolarmente soggiornanti in Italia. Noi Italia. Retrieved from https://noi-italia.istat.it/pagina.php?L=0&categoria=4&dove=ITA#:~:text=I%20cittadini%20non%20comunitari%20regolarmente,il%20doppio%20(%2B86%25).
- Italy Ministry of Enterprises and Made in Italy. (2012). Art. 25, Decreto-Legge 18 ottobre 2012, n. 179. Retrieved from https://www.mimit.gov.it/images/stories/Art25-dl179-2012.pdf
- Kovaleva Y., Hyrynsalmi S., Saltan A., Happonen A., Kasurinen J. (2023). Becoming an entrepreneur: A study of factors with women from the tech sector. Information and Software Technology 155.

- Kuziemko, I., Pan, J., Shen, Y., & Washington, E. (2018). The returns to working long hours. The Quarterly Journal of Economics, 133(1), 187-234.
- Ministry of Enterprises and Made in Italy. (2022). Quarto trimestre 2022: Relazione trimestrale sull’andamento economico. Retrieved from https://www.mimit.gov.it/images/stories/documenti/4_trimestre_2022.pdf
- Ministry of Enterprises and Made in Italy. (2024). Relazione annuale sullo stato di attuazione del Documento di Economia e Finanza (DEF) 2023. Retrieved from https://www.mimit.gov.it/images/stories/documenti/20240119_-_Relazione_annuale_DEF.pdf
- Ministry of Enterprises and Made in Italy. (2024). Secondo trimestre 2024: Relazione trimestrale sull’andamento economico. Retrieved from https://www.mimit.gov.it/images/stories/documenti/2_trimestre_2024.pdf
- Moore, G. A. (1993). The adoption of innovations: A study of the factors influencing the adoption and diffusion of software engineering innovations.
- OECD. (2007). Eurostat – OECD Manual on Business Demography Statistics, European Commission, Luxembourg.
- OECD. (2013), “Unemployment rates”, in OECD Factbook 2013: Economic, Environmental and Social Statistics, OECD Publishing, Paris.
- OECD. (2017), *Entrepreneurship at a Glance 2017*, OECD Publishing, Paris, https://doi.org/10.1787/entrepreneur_aag-2017-en.
- OECD. (2021), Entrepreneurship Policies through a Gender Lens, OECD Studies on SMEs and Entrepreneurship, OECD Publishing.
- Rey-Martí A., Tur Porcar A., Mas-Tur A. (2015). Linking female entrepreneurs' motivation to business survival. Journal of Business Research 68, 810-814.
- Skawińska, E., & Zalewski, R. I. (2020). Success factors of startups in the EU—A comparative study. Journal of Business Research, 12(10), 1-15.
- Stam E. (2009). Why Butterflies Don‘t Leave: Locational Behavior of Entrepreneurial Firms. Economic Geography 83, 27-50.
- tech.eu. (2023). The Global Startup Ecosystem Report 2023 shows how a year of growth and resilience in Europe. Retrieved from https://tech.eu/2023/06/15/the-

global-startup-ecosystem-report-2023-shows-how-a-year-of-growth-and-resilience-in-europe/

- Welter, F., Brush, C., & De Bruin, A. (2009). Gender embeddedness of women entrepreneurs: An empirical test of the 5M framework. ResearchGate.
- World Economic Forum. (2024). Global Gender Gap 2024 – Insight Report 2024.