Politecnico di Torino

Master of Science in Engineering and Management

The effect of Foreign Direct Investments on the emergence of ICT specializations in Italian provinces

Academic Year 2021/2022

March 2022

Supervisor:

Anna D'Ambrosio

Co-Supervisor:

Luigi Benfratello

Candidate:

Marcello Russo

Abstract

The analysis carried out in this thesis is aimed at checking whether the emergence of ICT specializations in Italian provinces is affected o not (and in case it is, to which extent) by an injection of capital from abroad through Foreign Direct Investments (FDIs).

Firstly, information about Italian companies from 1999 to 2019 has been extracted from AIDA dataset. Thereafter, all the extracted data have been re-organized in a more compact structure that includes the variables of interest (like the initials of the provinces' name or the timespan considered) for companies working in the ICT sector only.

Secondly, the content of fDi Markets dataset has been analyzed and filtered in order to highlight Foreign Direct Investments directed in Italy in ICT sector, from 2003 to 2019. Then, again the whole dataset has been re-organized such that it matches the structure used in the previous dataset for Italian companies.

Finally, the two datasets have been combined to create a single dataset containing all the information needed to carry on the analysis, linking the dependent variable (related to the specialization in ICT in a single province) to the independent ones (like the FDIs). Moreover, the dataset containing the number of ICT patents per Italian province has been re-organized and added to the final one, and so the number of patents is one of the independent variables as well.

The linear regression has been used for understanding the effect of independent variables on the ICT specialization, and more specifically the regression with binary dependent variables' theory has been applied (since the dependent variable here is a dummy).

General results from another paper lead to acknowledge a non-significant effect of FDIs on industry specialization, while they are actually effective in case there is already a technological identity in the region where the capital is injected. This effect is even reinforced in case of involvement of R&D activities, and it is also negatively related to the prior existing knowledge of the region. Therefore, this thesis wants to analyze the global effect of Foreign Direct Investments on local ICT specialization and verify whether the results are comparable with previous studies' ones or not.

Summary

1. Introduction 10
2. Background Literature 12
2.1 Foreign Direct Investments: definition, determinants and effects 12
2.1.1 Advantages and disadvantages of FDIs 16
2.2 Italian provinces: examples of specializations 17
2.3 ICT industry sector 20
2.4 FDIs' effects on specialization 22
3. Empirical Methods 26
3.1 Linear Regression 26
3.2 Regression with binary dependent variables 30
3.2.1 Probit Regression 31
3.2.2 Logit Regression 32
3.2.3 Maximum Likelihood Estimation 33
3.2.4 Measures of Fit 34
3.3 Empirical Model 35
4. Empirical Applications 36
4.1 Data Collection 36
4.1.1 Database AIDA (1999-2019) 37
4.1.2 Database FDIs (2003-2019) 38
4.1.3 Final database 40

4.2 Variables 41

4.2.1 Dependent variable 41

4.2.2 Independent variables 42

4.2.3 Control variables 42

4.3 Descriptive Statistics 43

5. Results 55

5.1 Average Marginal Effects 58

6. Conclusions 62

7. References 64

8. Appendix 68

Index of Figures and Tables

Figure 1: Horizontal vs Vertical FDI. 13
Figure 2: UNCTAD's Classification of FDI determinants (UNCTAD, 2002). 14
Figure 3: Italian subdivision among the three economic sectors (2012). 18
Figure 4: Example of industry specializations among municipalities (Komninos, 2005). 19
Figure 5: Employment level from 2000 to 2021 and 2025 employment projection for industry areas related to ICT. 21
Figure 6: Frequencies of observations in subperiods. 43
Figure 7: Frequencies of observations in Italian provinces. 44
Figure 8: Variable ict_capital summarized. 45
Figure 9: Variable comple_capital summarized. 45
Figure 10: Variable patents_pre summarized. 46
Figure 11: Variable ict_patents summarized. 46
Figure 12: Variable spec summarized. 47
Figure 13: Variable specICT summarized. 48
Figure 14: Most targeted Italian provinces by FDIs. 49
Figure 15: Number of ICT-related patents in Italian provinces. 50
Figure 16: Number of ICT-related FDIs from 2003 to 2006. 51
Figure 17: Number of ICT-related FDIs from 2007 to 2010. 51
Figure 18: Number of ICT-related FDIs from 2011 to 2014. 52
Figure 19: Number of ICT-related FDIs from 2015 to 2019. 53
Figure 20: Evolution of ICT specialization in Italian provinces (2003-2019). 54
Figure 21: Results of the probit regression model (first analysis). 55
Figure 22: Results of the probit regression model (second analysis). 57
Figure 23: Average marginal effects (first analysis). 59

Figure 24: Average marginal effects (second analysis). 60
Table 1: Cumulative Standard Normal Distribution. 68
Table 2: Cumulative Standard Normal Distribution (continued). 69
Table 3: Extract of database AIDA (1999-2019). 70
Table 4: Extract of database FDI (2003-2019). 71
Table 5: Extract of the final database. 72

1. Introduction

Multinational enterprises (MNEs) are often considered to be one of the primary entities to influence the globalization process through the integration of production processes across national boundaries by the transfer of capital and technology.
MNEs usually expand their activities in foreign countries for several reasons, such as exploitation of economies of scale, the use of specific advantage or just because their competitors are engaged in similar activities. Obviously, different countries, and so different economies might have policies altering things like corporate taxes, labor market conditions, subsides and so on, in order to be more attractive for injections of capital from abroad. One of the ways in which MNEs usually expand their operations in foreign countries is through Foreign Direct Investments (FDIs).

The term "FDIs" refers to investments that are made to acquire a lasting interest in an enterprise operating abroad. In other words, FDIs are international financial flows aimed at controlling or participating in the management of an enterprise in a foreign country. Since the late 1980s, global flows of FDIs have been increasing significantly: for many decades the majority of FDI flows have gone to developed economies. However, during the recent years, the share of FDI flows directed to developing and transition economies has increased. Many academic articles argue that FDIs can have important and positive effects on a host country's development effort. In addition to the direct capital financing they supply, FDIs can be a source of valuable technology and knowhow while encouraging linkages with local firms, which can give an economy a further growth push. Therefore, developing countries, as well as emerging economies, are often based on the assumption that greater inflows of FDIs will bring certain benefits to their economy, as FDIs are considered a key factor of modernization and economic development.

The objective of this thesis paper is to check whether FDIs actually bring benefits to the host country, in terms of industry specialization, or not. More specifically, the analyzed scenario wants to estimate the effect that ICT-related FDIs have on the emergence of new technological (ICT) specialization among Italian provinces. Therefore, the host country will be Italy, the analyzed industry sector will be the ICT sector and the timespan considered will be a 20-year period, from 1999 to 2019.

The thesis paper is organized as follows: Chapter 2 will give an overview of the main definitions and it will discuss a bit of the background literature on FDIs. Chapter 3 will provide the reader all the information about the linear regression and the probit regression model that have been used during the analysis, both theoretically and practically (describing the actual parameters of the probit regression model). Chapter 4 will describe what have been done in order to obtain the final database on which the analysis is based, what are the variables included in the analysis and it will provide a set of descriptive statistics. Finally, the results of the analysis are shown and commented in Chapter 5.

2. Background Literature

In this section, a bunch of definitions will be given at first, in order to introduce the reader into the context of foreign direct investments, industry specialization and ICT.
Then, an outlook of the studies made in the previous years will be provided, studies related to sectors different from ICT. The analysis carried out in this thesis will start from those results and is aimed at verifying the impact of foreign direct investments on ICT specialization in the Italian provinces, and if the obtained results are comparable with those related to other industry sectors.

2.1 Foreign Direct Investments: definition, determinants and effects

Investments (whether public or private, domestic or foreign) are crucial to the socio-economic transformation of any economy (Asongu, Akpan, & Isihak, 2018).
The whole thesis has been carried out focusing on the effect that foreign direct investments have on localized industry specialization, and so it is important to understand what FDIs are first, putting the attention on what the drivers that make foreign firms to invest are.

Foreign direct investment (FDI) is an investment from a party in one country into a business or corporation in another country with the intention of establishing a “lasting interest”.
An investment into a foreign firm can be considered an FDI only in the case it establishes a “lasting” interest for investors. To this end, investors have to obtain a minimum percentage of voting rights in the firm they are investing in (usually 10%).
Despite the lasting interest is an essential element for an investment to be defined as FDI, what really differentiates FDIs to passive foreign portfolio investments is the element of

control: investors want to be included in foreign firm's management decisions and operations, having the possibility to influence them (while in foreign portfolio investments they passively hold securities from a foreign country). This is the reason why the minimum percentage of voting rights in the foreign company is necessary to define FDIs.

It is generally possible to distinguish between horizontal and vertical FDIs. On the one hand, when a company aims at expanding its domestic operations to a foreign country, we are referring to a horizontal FDI (the company keeps doing the same activities as before, but now abroad as well). On the other hand, when a business wants to expand its activities to a foreign country by targeting a different level of the supply chain, we are referring to vertical FDI (the company does activities abroad that are different from those carried on domestically, but they are still related to the main business).

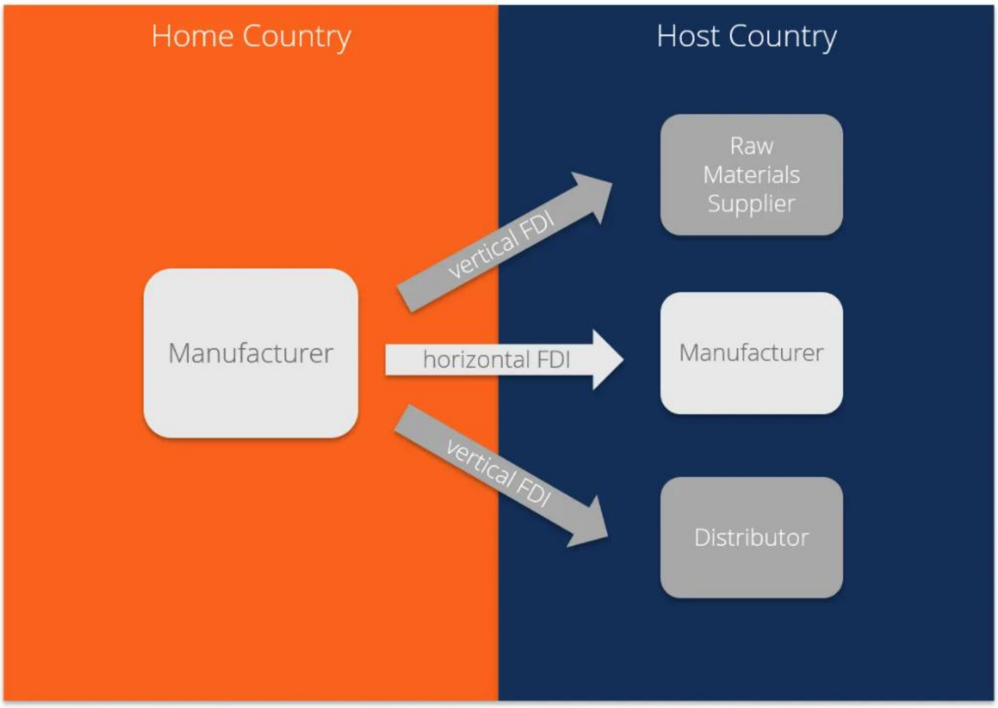

Figure 1: Horizontal vs Vertical FDI.

Finally, it is possible to define platform FDI as a foreign direct investment from a source country into a destination country for the purpose of exporting to a third country.

Once foreign direct investments have been defined, it is important to understand what the fundamental determinants of FDIs are. Since there are several findings and results, five particular determinants in the literature stand out. These are economic growth, market size, human capital, financial market development and infrastructure (Meivitawanli, 2021).

More generally, there are fundamental determinants of FDIs that are acknowledged by all of the versions of the contending theories, notably policy indicators (e.g., tax, trade, privatization, and macroeconomic policies), business dynamics (e.g., incentives for investment), market-related factors (e.g., market structure, market growth, and market size), resource-oriented determinants (e.g., technology availability, labor costs, and raw materials), and drivers toward economic efficiency (e.g., labor productivity, and transportation and communication costs) (Asongu, Akpan, & Isihak, 2018).

Determining Variables	Examples
Policy variables	Tax policy, trade policy, privatization policy, macroeconomic policy
Business variables	Investment incentives
Market-related economic determinants	Market size, market growth, market structure
Resource-related economic determinants	Raw materials, labor cost, technology
Efficiency-related economic determinants	Transport and communication costs, labor productivity

Figure 2: UNCTAD's Classification of FDI determinants (UNCTAD, 2002).

The reason why firms engage in FDIs is related to the existence of specific assets whose value is higher under foreign control, which allows firms to compete in foreign environments. This view credits the genesis of FDIs to the possession of some assets, such as technology or know-how, that constitutes a significant gain for the host country.

This, in turn, suggests that FDIs can play an important role in accelerating and modernizing a country's economic growth (Alfaro, 2016). The general effects of Foreign Direct Investments are various and depend on different aspects, literature says (Markusen & Venables, 1999) (Görg & Greenaway, 2004).

FDIs are an important vehicle for transferring technology and promoting growth only when the host country has a minimum threshold of human capital (Borensztein, Gregorio, & Lee, 1998). However, most developing countries do not meet such a threshold (Xu, 2000).

Moreover, underdevelopment of local financial markets can limit an economy's ability to exploit the potential of FDI spillovers (Alfaro, Chanda, Kalemli-Ozcan, & Sayek, 2004). The results of the analysis of growth on FDI to GDP, together with various controls variables, indicate that FDI, on its own, does not exert a robust positive impact on growth. When the interaction term is included, however, the regression results become positive and significant, leading to the result that the positive benefits of FDI is contingent a country's possession of a strong financial sector.

The relation between FDI and growth turns out to be stronger for industries that rely more on external financing (Alfaro & Charlton, 2013) and, obviously, host countries with more developed financial markets attract more multinational entry (Bilir, Chor, & Manova, 2014).

Another distinction that arises is in the effect on a country when considering horizontal or vertical FDIs (Alfaro, 2016):

- horizontal FDIs may raise income in each country without necessarily changing its distribution;
- vertical FDIs may reduce absolute wage differences across countries and alter relative wages within countries.

Finally, among the effects of Foreign Direct Investments, the one that drives this analysis is the development of technological specializations in the host country/region following an injection of capital from abroad. In particular, the effect of FDIs with respect to the development of a regional specialization is formalized in the following paragraph (2.1.4).

2.1.1 Advantages and disadvantages of FDIs

What people may generally think is that both the investor and the foreign host country benefit from FDIs, and so they usually have incentives to allow them. In reality, things are a little bit more complicated than that. Several studies have tackled this topic with the ultimate goal of defining whether Foreign Direct Investments are positive or not for both the investor and the foreign host country, but the results are still not so straightforward (Alfaro, 2016). What is clear is instead that Foreign Direct Investments provide advantages and disadvantages to both companies and the host country, and they are summarized as follows (Szanyi, 1998) (Obalade, 2014) (Bose, 2012) (Cuervo-Cazurra, 2017).

Some of the advantages of FDIs for companies are:

- Market diversification;
- Tax incentives;
- Lower labor costs;
- Preferential tariffs;
- Subsidies.

The following are instead some of the advantages for the host country:

- Economic stimulation;
- Development of human capital;
- Increase in employment;

- Access to management expertise, skills, and technology.

Companies' advantages are mostly related to cost-cutting and lowering risk, while for host countries they are mainly economic.
However, there are still two main drawbacks to FDI:

- Displacement of local businesses
- Profit repatriation

The entry of large firms in another country may displace local businesses. In the case of profit repatriation, the drawback is related to the fact that firms will not reinvest profits back into the host country, and this leads to large outflow of capital from the host country. The consequent result then, is that many countries have regulations limiting foreign direct investment.

2.2 Italian provinces: examples of specializations

The expression "industry specialization" refers to a series of strategies that companies follow to maximize productivity, knowledge and leadership in the targeted field. To this end, they aim at focalizing the whole business in the production of products and/or services related to that sector.
Specialization may also refer to provinces of a region, regions of a country or even whole nations. Several countries around the world specialize in producing goods or delivering services that are native to their geographical area, and they import other goods and services.

Italy is a country made of 20 regions and 107 provinces. First of all, it is possible to identify a subdivision, among municipalities of provinces, related to the economic sector they are more specialized in.

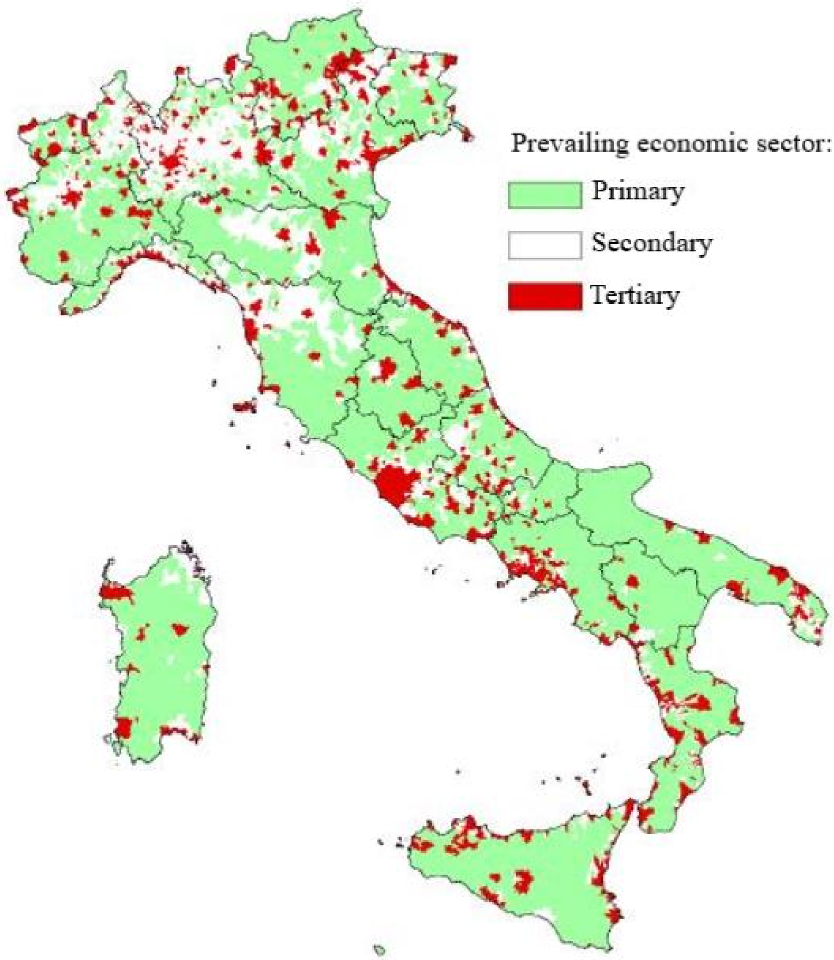

Figure 3: Italian subdivision among the three economic sectors (2012).

What comes out from Figure 3 is that more than a half of Italy is specialized in primary sector activities (58.7%), followed by secondary sector activities that are very diffused as well (31.4%). Finally, tertiary sector activities occupy only a little part in the Italian economic sectors' subdivision with the low percentage of 9.9%.

Having discussed about the economic sectors' subdivision, in Figure 4 (below) there is an example of industry specialization among the Italian municipalities in order to be a bit more specific on which kind of sectors are predominant among the whole country.

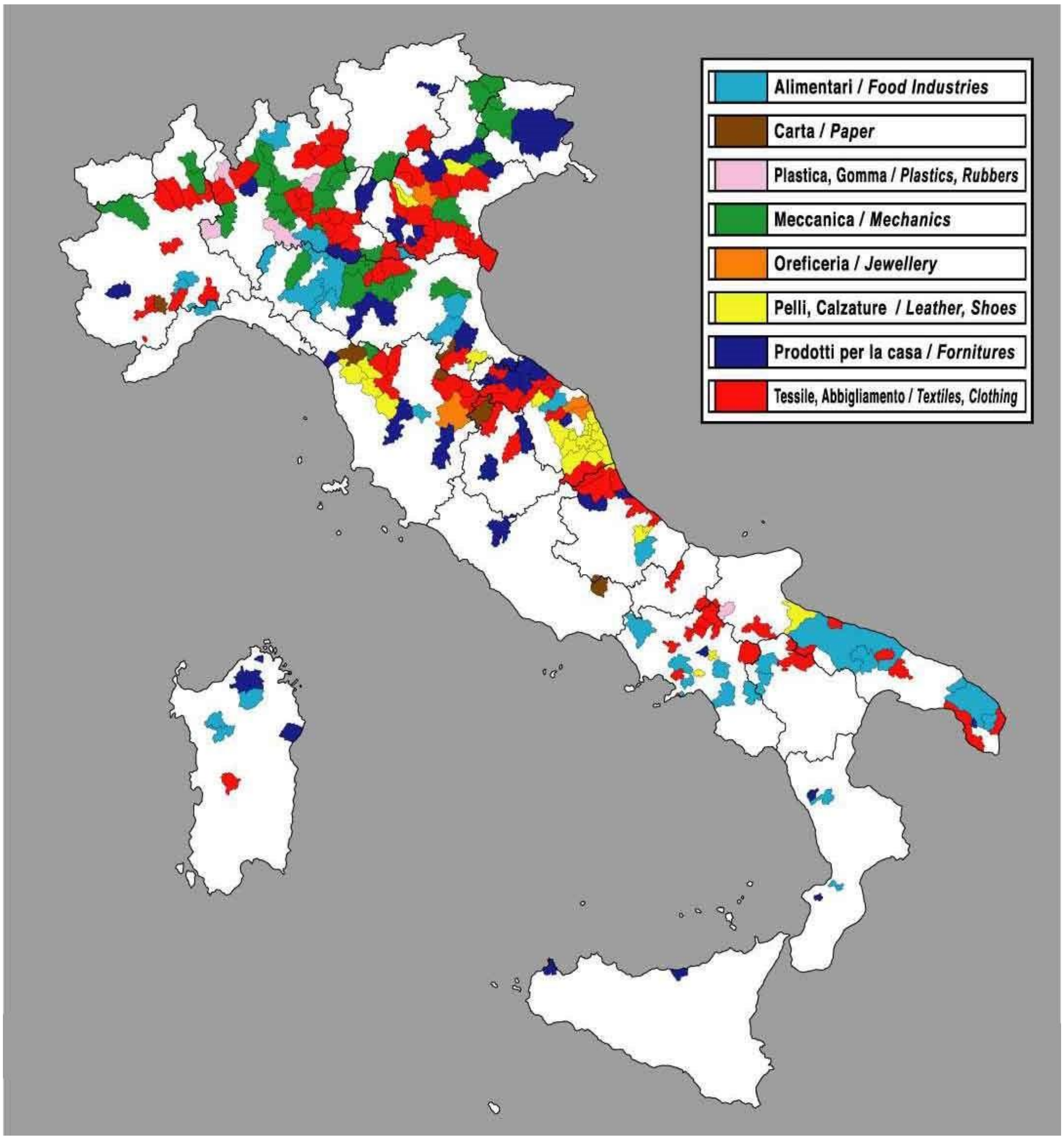

Figure 4: Example of industry specializations among municipalities (Komninos, 2005).

2.3 ICT industry sector

The term "ICT" stands for Information and Communication Technology and commit to both communication networks and the various technologies used in them.

ICT sector refers to equipment and services related to broadcasting, computing and telecommunications, all of which capture and display information electronically (2004). The contribution of this sector to technological progress and productivity growth is great, and its impact can be examined in two ways:

- directly, focusing on how much it contributes to output, employment or productivity growth;
- indirectly, as a driver of technological change influencing other parts of the economy.

The Information and Communications Technology (ICT) industry can be subdivided into four main areas:

- telecommunications services;
- internet service providers, web search portals and data processing services;
- computer system design and related services;
- internet publishing and broadcasting.

The following chart represents the employment level in each of the areas mentioned above, from 2000 to 2021 and a forecast of that index directly to 2025.

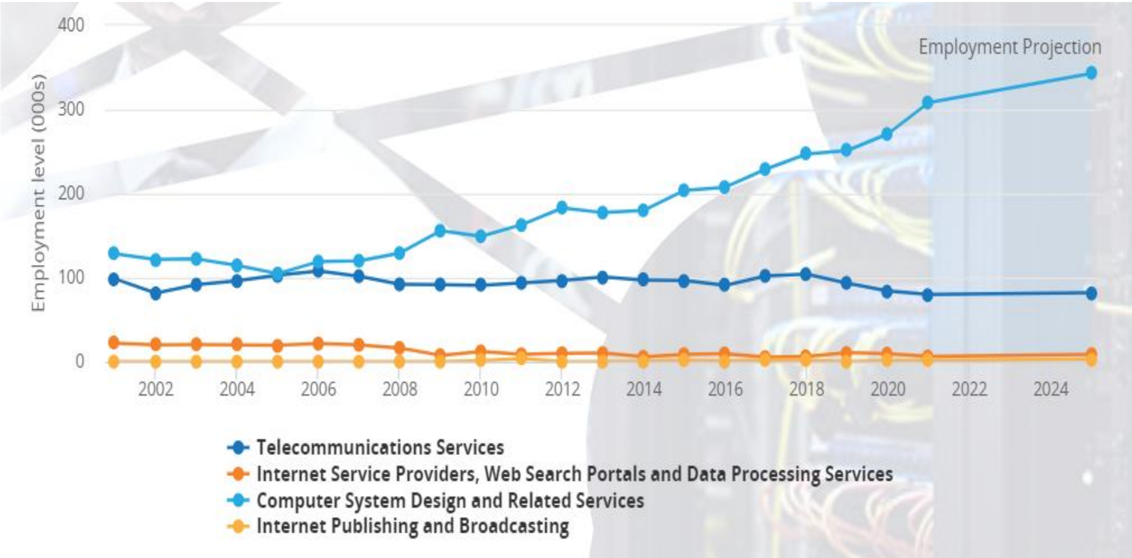

Figure 5: Employment level from 2000 to 2021 and 2025 employment projection for industry areas related to ICT.

Analyzing the output of the chart in Figure 5, it is possible to state that the largest sector by far is Computer System Design and Related Services. This sector has seen strong growth over the past couple of decades, with the employment level increasing by nearly 14% between 2020 and 2021. Moreover, the employment level is projected to reach a value close to 350,000 by 2025.

The second largest sector is Telecommunications Services. Despite the trend between 2001 and 2021 has been downward sloping, the projection to 2025 shows a slight increase in employment levels to about 82,000.

Employment levels in the Internet Service Providers, Web Search Portals and Data Processing Services sector have dropped significantly from 2001 to 2021, although there was an increase in 2019. The level is projected to increase to about 9,000 in 2025.

2.4 FDIs' effects on specialization

The literature on the effect of FDIs considers several different situations in which the presence or the absence of one or more variables can have a higher or lower influence on the emergence of a new technological specialization.

The first general conclusion on the effect of FDIs that comes out from the studies made by Castellani, Marin, Montresor and Zanfei. In 2020, they carried out the analysis of the effect of FDIs on regional specialization in environmental (green) technologies, and the first outcome of that is the following: FDIs as a whole have non-significant impact on regional specialization (Castellani, Marin, Montresor, & Zanfei, 2020). However, this is just the first general result they obtained.

What can really makes the difference is an already established technological identity in the place where capital injections are directed. This is the case in which FDIs turn out to be positive and significant on the specialization process of the region in which they are injected.

Moreover, their effect is even stronger if they involve R&D activities. The influence of R&D activities is very powerful since they increase the knowledge of that place directly and favor the occurrence of technological improvements.

As a consequence, FDIs in R&D activities allow regions to keep being specialized in a technological sector in time. This is true as long as they were already specialized in that, since if it is not, FDIs in R&D activities do not facilitate the switch from non-specialized to specialized region. In particular, this result is valid for average levels of relatedness of the new technologies to the pre-existing specializations of the region. On the contrary, for high levels of unrelatedness between new technologies and the previous existing technological specialization, FDIs in R&D activities can positively impact on the regions' switch.

The relatedness to already mastered technologies is something that, according to recent developments in the geography of innovation (Balland, 2016), is expected to drive regions' capacity to specialize and diversify into a specific technological domain. This variable is meant as a synthetic measure of the cognitive proximity of the former to the latter (Boschma, 2015).

The literature on technological diversification has shown that many technologies develop in a path- and place-dependent way, conditionally on the existing (regional) knowledgebase (Berge & Weterings, 2014) (Tanner, 2016) (Barbieri, Perruchas, & Consoli, Specialization, diversification and environmental technology-life cycle, 2018) (Colombelli & Quatraro, 2019) (Corradini, 2019) (Barbieri & Consoli, 2019) (Montresor & Quatraro, 2019) (Consoli, Castellacci, & Santoahla, 2019) (Santoalha & Boschma, 2019).

Nevertheless, it is quite difficult to accurately forecast whether inward FDIs increase or not the knowledge base of places and affect their technological specialization (Castellani, Marin, Montresor, & Zanfei, 2020). This is mainly due to the combination of heterogeneous local Multi National Enterprises (MNEs) strategies (Marchi, Maria, Khrishnan, & Ponte, 2020) and region-specific regulation stringency and technological competencies/capabilities (Montresor & Quatraro, 2019).

Studies have also shown that foreign firms' activities in specific technological domain can contribute, indirectly, to increase the identity of domestic firms in that same domain (Albornoz, Cole, Elliott, & Ercolani, 2009) (Dechezlepretre & Glachant, 2014) (Cainelli, Mazzanti, & Montresor, 2012), but still depending on a set of circumstances (Rezza, 2013) (Tang, 2015).

An additional aspect that has to be analyzed is related to the functional activities through which MNEs can affect the technological specialization of regions. They are those activities related to Research and Development and innovation.

R&D FDIs are likely to provide both higher direct contribution to local innovation and a potential for significant spillovers on the innovation of local firms (Braconier, Ekholm, & Knarvik, 2001) (Castellani & Zanfei, 2006) (Fu, 2008) (Marin & Sasidharan, 2010) (Todo, 2006) (Belitz & Molders, 2016).

Concluding the discussion above, it is possible to say that the effects of inward FDIs on regional technological specialization are quite likely to depend on the nature of such activities both across industries and across functional domains.

Another aspect that has to be considered during the analysis is the capacity of regions to diversify their technological identity over time.

In regions with an already established specialization, inward FDIs (especially those related to the specialization) can inject additional knowledge and competencies to maintain that level of specialization over time, or even reinforce it. In fact, a region's capacity to maintain a specialization could diminish over time. The absorption of external knowledge and experience through FDIs could reduce the risk of an "inverse transition", from specialization to non-specialization (Castellani, Marin, Montresor, & Zanfei, 2020). Moreover, there may be reason to believe that FDIs will also help regions gain a new tech advantage from scratch, should they not have it already, but the actual contribution of FDIs in acquiring a specialization from scratch, or in keeping an existing one is something with respect to which literature does not provide an empirical answer *a priori*.

Finally, it is possible to state that the relatedness of pre-existing technologies to the new ones in the regional knowledgebase is something that for sure can favor the regional specialization in that specific technological field.

Considering some previous studies made, factors that can (positively or negatively) influence the effect of relatedness on technological specialization have been identified.

On the one hand, FDIs bring to the hosting region external knowledge and competencies, which make the development of specific technologies less place dependent.

On the other hand, the specific technological content of inward FDIs could overlap with the actual regional knowledge base and reinforce previous specialization patterns.

3. Empirical Methods

3.1 Linear Regression

Once all variables have been determined and explained, and once the empirical correlation between specialization and foreign direct investments has been analyzed, the validity of the results of the descriptive statistics has to be checked.
The impact that foreign direct investments have on the emergence of a new ICT specialization has been estimated by the parameters of the linear regression model and so, a brief overview of the linear regression itself will be given at first.

Linear regression represents a method of estimation of the linear relationship between two variables. To this end, the objective is to determine the value of the slope parameter of the population regression line, which is the expected effect on Y of a unit of change in X (i.e., $\frac{\Delta Y}{\Delta X}$).
The general notation of the population regression line is the following:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \quad i = 1, \dots, n \tag{3.1}$$

where:

- X is the independent variable of "regressor";
- Y is the dependent variable of "regressand";
- β_0 is the intercept of the population regression line;
- β_1 is the slope of the population regression line;
- u_i is called "regression error" or simply "error".

While the intercept and the slope of the population regression line are not known and must be estimated using sample data, the error term represents unobserved variables that still affect Y but are different from X.

One way of estimating the two parameters β_0 and β_1 is by using the Ordinary Least Squares (OLS) estimator. It minimizes the average squared difference between the actual values of Y and the prediction based on the estimated line, the "predicted value".
Therefore, the values of the slope and the intercept of the population regression line, estimated by using the OLS estimator, are respectively:

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \tag{3.2}$$

$$\widehat{\beta_0} = \bar{Y} - \widehat{\beta_1}\bar{X} \tag{3.3}$$

Whereas the OLS predicted values and residuals are respectively:

$$\widehat{Y_i} = \widehat{\beta_0} + \widehat{\beta_1}X_i, \quad i = 1, \dots, n \tag{3.4}$$

$$\widehat{u_i} = Y_i - \widehat{Y_i}, \quad i = 1, \dots, n \tag{3.5}$$

Something that is notable to mention is that there might be situations in which the value of the intercept is different from zero, but it can be meaningless (it makes no sense in practice). In that case, it is said that the intercept has geometrical interpretation only.

The following step is to check for the goodness of fitting of the estimated linear regression line (that is, whether it represents data in a proper way or not). To this end, two complementary statistics are used:

- The coefficient of determination R^2, that measures the fraction of variance of Y that is explained by X; it is unitless and it ranges between 0 (no fit) and 1 (perfect fit);
- The standard error of the regression SER, that measures the magnitude of a typical regression residual in the units of Y (it measures the dispersion of the distribution of u).

Firstly, the coefficient of determination is expressed as follow:

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{3.6}$$

On the one hand, when the variable X does not explain the variance of Y, then the coefficient of determination will be equal to zero ($R^2 = 0$ that means $ESS = 0$). On the other hand, when the variable X explains all of the variance of Y, the coefficient of determination will be equal to one ($R^2 = 1$ that means $ESS = TSS$ and $Y = \hat{Y}$).

Secondly, the standard error of the regression is computed as:

$$SER = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n} \hat{u_i}^2} \tag{3.7}$$

The division by $n - 2$ is a "degree of freedom" correction related to the fact that for the *SER*, two parameters have been estimated (β_0 and β_1, by $\widehat{\beta_0}$ and $\widehat{\beta_1}$ respectively).

Moreover, another way to check for the goodness of fitting is that of using the root mean squared error (*RMSE*). It is closely related to *SER*, and it is expressed as follow:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \hat{u_i}^2} \tag{3.8}$$

This measures the same thing as the *SER*, with the only difference in the division by $1/n$ instead of $1/n - 2$.

Apart from the general scenario presented before, there can be situations in which a regressor is binary. When it is the case, the only admissible values that the regressor can assume are 0 and 1. Moreover, binary regressors are usually called "dummy" variables and it is notable to say that it does not make sense to call β_1 as the "slope" of the regression line if the regressor is a dummy variable.

Considering the general expression of the population regression line (3.1), when X is binary:

- when $X_i = 0$, then $Y_i = \beta_0 + u_i$ (the expected value of Y_i is β_0);

- when $X_i = 1$, then $Y_i = \beta_0 + \beta_1 + u_i$ (the expected value of Y_i is $\beta_0 + \beta_1$).

Finally, let now us consider the case in which there is more than one regressor (multiple regression). The expression (3.1), considering two regressors, becomes:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i, \quad i = 1, \dots, n \tag{3.9}$$

where:

- Y is the dependent variable;
- X_1 and X_2 are the two independent variables or regressors;
- β_0 is the unknown population intercept;
- β_1 is the effect on Y of a change in X_1, holding X_2 constant;
- β_2 is the effect on Y of a change in X_2, holding X_1 constant;
- u_i is the regression error (related to omitted factors that are different from X_1 and X_2 but still affect Y).

In case of multiple regression, there are two types of measures of the fit:

- Those based on residuals size, that are the standard deviation of $\widehat{u_i}$ with (*SER*) and without (*RMSE*) degrees-of-freedom correction;
- Those based on the explained variance fraction, that are the fraction of variance of Y explained by X (R^2) and the "R^2 adjusted" ($\overline{R^2}$), that is the R^2 with a degrees-of-freedom correction for estimation uncertainty ($\overline{R^2} < R^2$).

The two measures based on residuals size are:

$$SER = \sqrt{\frac{1}{n-k-1} \sum_{i=1}^{n} \widehat{u_i}^2} \tag{3.10}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \widehat{u_i}^2} \tag{3.11}$$

Whereas the two measures based on the explained variance fraction are:

$$R^2 = \frac{ESS}{TSS} \quad (3.12)$$

$$\overline{R^2} = 1 - (\frac{n-1}{n-k-1})\frac{SSR}{TSS} \quad (3.13)$$

with $SSR = \sum_{i=1}^{n} \widehat{u_i}^2$ and *TSS* that is the same as before, expressed in (3.6).

The addition of another regressor never makes R^2 decrease and so this is a problem, being R^2 a measure of fit. For this reason, $\overline{R^2}$ has been introduced. It corrects this problem by giving a sort of "penalization" for increasing the number of regressors.

3.2 Regression with binary dependent variables

So far, the independent variables of the linear regression model have been considered either binary or non-binary, with no particular issues in both cases. But what happens when the dependent variable is binary? This is the case here, since the dependent variable aims at highlighting the presence of a local ICT specialization or not. Therefore, the only admissible values are 0 and 1. In this case, things are obviously more difficult since the regression function has to be interpret as a predicted probability (Stock & Watson, 2014).

The linear multiple regression model applied to a binary dependent variable is called the linear probability model ("linear" because it is a straight line and "probability" model because it models the probability that the dependent variable equals 1).

Within this model, the population coefficient β_1 related to the regressor *X* represents the change in the probability that $Y = 1$ associated to a unit change in X_1, holding the other regressors constant (and so on for $i = 1, \ldots, k$). The same changes apply to the estimated predicted values.

The linear probability model is the linear multiple regression model expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + u_i \quad (3.14)$$

Since *Y* is binary, we have that:

$$\Pr(Y = 1|X_1, \dots, X_k) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \tag{3.15}$$

The regression coefficients can be estimated by OLS and, apart from R^2, the OLS standard errors can still be used. The problem with R^2 is the following: while for continuous dependent variables it can equal 1 (all data lie exactly on the regression line), this is impossible for binary dependent variables (unless the regressors are binary as well). One of the main issues related to a binary dependent variable is related to the fact that, since probabilities range between 0 and 1, the effect on the probability that $Y = 1$ of a given change in X must be *non-linear*. Very high (low) values of X might lead to values of the binary dependent variable higher (lower) than one (than zero): this is totally nonsense since we are talking about a probability. To address this problem, the probit and logit regression models have been introduced.

Probit and logit regression are nonlinear regression models specifically designed for binary dependent variables, since they force predicted values to be between 0 and 1 using cumulative probability distributions (c.d.f.'s). Probit regression uses the standard normal cumulative probability distribution, whereas logit regression (also called logistic regression) uses the "logistic" cumulative probability distribution.

3.2.1 Probit Regression

The probit regression model with a single regressor X is:

$$\Pr(Y = 1|X) = \Phi(\beta_0 + \beta_1 X) \tag{3.16}$$

Where Φ is the cumulative standard normal distribution function (tabulated in Appendix Table 1 and Table 2).

In the probit model, the term $\beta_0 + \beta_1 X$ is the so-called "z" in the cumulative standard normal distribution (Table 1 and Table 2). The probit coefficient β_1 in the equation (3.16) is the change in the z-value associated to a unit change in X. Considering:

- $\beta_1 > 0$: an increase in *X* increases the value of z and thus the probability that $Y = 1$ increases;
- $\beta_1 < 0$; an increase in *X* decreases the probability that $Y = 1$.

It is notable to say that, although the effect of *X* on the *z*-value is linear, its effect on the probability is nonlinear.

Considering now not one but two regressors (probit regression with multiple regressors), the equation (3.16) becomes:

$$\Pr(Y = 1|X_1, X_2) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2) \tag{3.17}$$

In the case in which the population regression function is a non-linear function of *X*, the expected change in *Y* due to a change in *X* is still estimated in three steps (being the expected change in *Y* due to a change in *X* the change in the probability that $Y = 1$). First of all, compute the predicted value of *Y* with the original value of *X*. Then, compute the predicted value of *Y* with the new value of *X*. Finally, compute the difference between the two predicted values of *Y*.
The general expression of the (3.17) with multiple regressors is the following:

$$\Pr(Y = 1|X_1, \dots, X_k) = \Phi(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k) \tag{3.18}$$

The probit coefficients are estimated by the method of maximum likelihood, which produces efficient estimators in a wide variety of applications. The maximum likelihood estimator in consistent and normally distributed in large samples, so *t*-statistics and confidence intervals for the coefficients can be constructed in the usual way.

3.2.2 Logit Regression

The logit regression model of the binary dependent variable *Y* with multiple regressors is:

$$\Pr(Y = 1|X_1, \ldots, X_k) = F(\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k)$$
$$= \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\cdots+\beta_k X_k)}} \quad (3.19)$$

Logit regression is similar to probit regression except for the cumulative distribution function, which is different from the cumulative standard normal distribution function. In particular, here the cumulative standard logistic distribution function is used, denoted by *F*. Its form is defined in terms of exponential function, which is given by the equation (3.19).

As with probit, the logit coefficients can be estimated by maximum likelihood. The maximum likelihood estimator is consistent and normally distributed in large samples, so t-statistics and confidence intervals for the coefficients can be constructed in the usual way.

The differences between probit and logit regression functions are small and they often produce similar results. Logit regression used to be preferred because of the faster computation of the logistic cumulative distribution function, but now with more efficient computers this difference is no longer important.

3.2.3 Maximum Likelihood Estimation

After having defined them, it is possible to state that the probit and logit regression functions are a non-linear function of the coefficients (i.e., the coefficients β_0, β_1,…, β_k appear *inside* the cumulative standard normal distribution function in case of probit regression, and *inside* the cumulative standard logistic distribution function in case of logit regression). For this reason, the coefficients β_0, β_1,…, β_k cannot be estimated by OLS, but they are instead estimated by maximum likelihood.

The likelihood function is the joint probability distribution of the data, treated as a function of the unknown coefficients. The maximum likelihood estimator (MLE) of the unknown coefficients consists of the values of the coefficients that maximize the likelihood function, which is in turn the joint probability distribution.

The likelihood function for $n = 2$ independent and identically distributed observations (Y_1, Y_2) on a binary dependent variable with no regressors is:

$$f(p;\ Y_1, Y_2) = p^{(Y_1+Y_2)}(1-p)^{2-(Y_1+Y_2)} \tag{3.20}$$

where the only unknown parameter to estimate is the probability p that $Y = 1$, which is also the mean on *Y*.

The maximum likelihood estimator of p is the value of p that maximizes the likelihood function in the equation (3.20). For general *n*, the MLE $\hat{p}$ of the Bernoulli probability p is the sample average (that is, $\hat{p} = \bar{Y}$).

3.2.4 Measures of Fit

As mentioned before, the R^2 is a poor measure of fit for the linear probability model, and this is also true for probit and logit regression. The two measures of fit used for models with binary dependent variables are the "fraction correctly predicted" and the "pseudo-R^2".

The fraction correctly predicted says that if $Y_i = 1$ and the predicted probability is higher than 50%, or if $Y_i = 0$ and the predicted probability is lower than 50%, then Y_i is correctly predicted. Otherwise, Y_i is said to be incorrectly predicted. The fraction correctly predicted is the fraction of the *n* observations $Y_1, \dots, Y_n$ that are correctly predicted.

The pseudo-R^2 measures the fit of the model using the likelihood function. Since the MLE maximizes the likelihood function, adding another regressor to the probit or logit model increases the value of the maximized likelihood. Therefore, the fit is measured by comparing values of the maximized likelihood function with all the regressors to the value of the likelihood with none.

Specifically, the pseudo-R^2 for the probit model is:

$$pseudo - R^2 = 1 - \frac{\ln(f_{probit}^{max})}{\ln(f_{Bernoulli}^{max})} \tag{3.21}$$

3.3 Empirical Model

The dependent variable considered for the analysis is a binary variable that equals 1 when there is an established ICT specialization in the province i at time t, and 0 if not. It is named $specICT_{it}^{0,1}$ and, it is obtained by considering the number of companies that are operating in the ICT sector, per Italian province. The next chapter will provide all the information needed about the variables and how they are obtained/computed.

The baseline specification is a probit estimation of the following model:

$$specICT_{it}^{0,1} = \Phi(\alpha + FDI_{it}\beta + X'_{it}\gamma + \lambda_t + \varepsilon_{it}) \tag{3.22}$$

where:

- Φ is the cumulative standard normal distribution function;
- FDI_{it} is the independent variable related to FDIs in ICT;
- X'_{it} is the vector of control variables (FDI in complementary sectors to ICT, already established ICT specialization and the number of ICT patents before the injection of foreign capital);
- λ_t is a series of period-specific dummies to account for time-varying unobserved features;
- ε_{it} is the error term with standard properties;
- α, β and γ are the coefficients of the probit regression model.

As already said, the definition of the single variables will be presented in the following chapter.

4. Empirical Applications

4.1 Data Collection

To begin with, the timespan considered for the analysis is a 20-year period, going from 1999 to 2019. The 20-year timespan has been then subdivided into five subperiods:

- from 1999 to 2002;
- from 2003 to 2006;
- from 2007 to 2010;
- from 2011 to 2014;
- from 2015 to 2019.

The first subperiod has been considered as a sort of baseline, since no Foreign Direct Investments are available for that period and so it can be considered as a starting point. Moreover, it might be interesting to consider the period from 2011 to 2014 since it is the first subperiod following the Italian crisis that took place in 2008 (during the third subperiod).

The final database is a combination of two other datasets: the first one, containing the number of companies operating in each Italian province during the considered period, and the second one related to the Foreign Direct Investments done in the past years.

The first dataset has been taken from the AIDA website, downloading information about Italian companies related to the period of interest. Information like the name of the company, where it operates, the net income, the number of employees during the years have been considered. AIDA is the database created and distributed by Bureau van Dijk S.p.A., containing the financial statements, the personal and product data of all the active and defaulted Italian companies (with the exception of Banks, Insurance Companies and Public Bodies).

The second dataset comes from the fDi Markets website. It contains all the Foreign Direct Investments made in several countries and several industry sectors among the years, specifying information like the destination city, the date and obviously the industry sector. fDi Markets is the most comprehensive online database of cross-border investments available, covering all countries and sectors worldwide. It is provided by the Financial Times, and it gives access to real-time monitoring of investment projects, capital investment and job creation. It is also possible to track and profile companies investing overseas, as well as conduct in-depth analysis to uncover trends.

4.1.1 Database AIDA (1999-2019)

The first database is made up of 10 variables on columns and 184,972 observations on rows. For each Italian province, and for each subperiod and each industry sector, it shows whether there was an ICT specialization or not in that province, and in that subperiod.

As shown in Table 3 in the Appendix, four variables "count" have been created in order to compute the specialization index. Here there are all the variables:

- "sigla" refers to the initials of the single Italian province to which the observation is referred;
- "subperiod" states to which subperiod among the ones above the observation is referred;
- "ateco_3d" indicates the first three digits of the ATECO code (2007) that identify the industry sector to which the observation is referred;
- "count_prov_sector" reports the number of firms operating in the industry sector indicated by the ATECO code, for the specified Italian province in that subperiod;
- "count_prov" reports the number of firms operating in the Italian province during the specified subperiod;

- "count_sector" reports the number of firms operating (in Italy) in the industry sector indicated by the ATECO code, during that subperiod;
- "count_tot" reports the total number of firms operating in Italy during the specified subperiod;
- "spec" reports the so-called "coefficient of localization", which is computed as in the Equation (4.1) and it gives the value related to the actual specialization of the province (the province is said to be specialized when spec is higher than 1) in the industry sector specified by the ATECO code;
- "ICT" is a dummy variable that checks whether the industry sector related to the observation is ICT or not (being the first three digits of the ATECO code for ICT equal to 620);
- "specICT" is the dummy, dependent variable considered for the analysis that equals 1 when "spec" is higher than 1 and "ICT" is equal to 1 (i.e., when there is an ICT specialization in the Italian province, in that subperiod), otherwise it is zero.

Lately, the variable "specICTl1" has been added. It reports whether there was an already established specialization in ICT in the Italian province or not in the subperiod immediately before the one considered. The "l" in the variable name stands for "lag".

4.1.2 Database FDIs (2003-2019)

The second database is made up of 19 variables on columns and 2901 observations on rows. It contains all the information about every Foreign Direct Investment in Italy, specifying details like the source state, the investing company, the Italian destination province, the industry sector and the capital invested. An extract of the database is shown in Table 4 in the Appendix, and here there are all the variables:

- "ProjectDate" refers to the date of the foreign investment;

- “InvestingCompany” reports the name of the foreign investing company;
- “ParentCompany” reports the name of the parent company of the investing one, in case they differ;
- “SourceCountry” reports the name of the source country;
- “SourceState” specifies the state of the source country from which the investment is made;
- “SourceCity” specifies the city of the source state from which the investment is made;
- “DestinationCountry” reports the destination country (Italy) of the Foreign Direct Investment;
- “DestinationState” specifies the Italian region in which the capital is injected;
- “AdminRegion” specifies the Italian province in which the capital is injected;
- “DestinationCity” specifies the city of the Italian province in which the capital is injected;
- “IndustrySector” reports the industry sector in which the Foreign Direct Investment is directed;
- “SubSector” specifies the sub-sector in which the capital is injected;
- “Cluster” reports the activities’ operation area in which the capital is injected;
- “IndustryActivity” specifies the activity operation where the investment is directed;
- “CapitalInvestment” reports the amount of capital invested;
- “Estimated” tells whether the value of invested capital is real or estimated;
- “JobsCreated” reports the number of additional jobs created following the investment;
- “S” tells whether the number of jobs created is real or estimated;

- “ProjectType” tells if the investment project is related to an expansion of an already existing project, or to a brand new one.

4.1.3 Final database

The final research sample is a combination of the two previous datasets, from which the main variables have been taken. It is made up of 44 variables on columns and 436 observations on rows. An extract of the final database is shown in Table 5 in the Appendix. Apart from the already defined variables deriving from the previous datasets, there are a few new variables (introduced as “control variables”) that will be better explained in the next paragraph. The most important ones are:

- “ict_patents” reports the number of patents in the field of ICT published in the specified subperiod;
- “patents_pre” reports the number of patents in the field of ICT in a specific Italian province related to the first subperiod, going from 1999 to 2002;
- “fdi_ict” is a variable that counts for the number of Foreign Direct Investments in the field of ICT, in a specific Italian province and in a specific subperiod;
- “pre_specICT” is the variable that reports whether there was already a specialization in ICT in a province in the first subperiod (1999-2002);
- “pre_spec” reports the so-called “coefficient of localization”, which is computed as in the Equation (4.1), and it gives the value related to the actual ICT specialization of the province (the province is said to be specialized when spec is higher than 1) in the first subperiod (1999-2002);
- “comple_fdi” is a dummy variable that equals 1 when there are Foreign Direct Investments in complementary sectors to ICT (in a specific Italian province and subperiod), and 0 if not.

Moreover, the industry sectors that have been considered as complementary to ICT are:

- electronic components;
- consumer electronics;
- semiconductors.

4.2 Variables

This section will provide the reader a brief overview of all the variables of interest used during the analysis. Firstly, it will be shown how the dependent variable is defined. Thereafter, the independent variables considered for the analysis will be presented. Finally, all the considered control variables will be introduced.

4.2.1 Dependent variable

Being the objective of this analysis the evaluation of the effect of Foreign Direct Investments on the emergence of ICT specializations, the dependent variables of the probit regression model is a dummy variable that equals 1 when the Italian province is said to be specialized in ICT during the specific subperiod, and 0 if not. The actual value of specialization comes from the variable "spec", that is computed as follow:

$$spec = \left(\frac{count\_prov\_sector}{count\_prov}\right) \Big/ \left(\frac{count\_sector}{count\_tot}\right) \tag{4.1}$$

The value of the variable "specICT" depends on the value of the variable "spec": when the value of "spec" is equal or higher than 1, then the value of "specICT" will be 1 and the related Italian province is said to be specialized in ICT in the considered subperiod. On the contrary, when "spec" is lower than 1, the value of "specICT" will be 0 and there is no ICT specialization in the province.

4.2.2 Independent variables

The independent variable of the probit regression model is, in general, the one that shows the presence (and in case, the number) or absence of ICT-related Foreign Direct Investments in Italian provinces. Therefore, the variable will be the one called "fdi_ict" presented before, that is defined in the final research sample shown in Table 5 in the Appendix (an extract of it).

Nevertheless, it may be useful to consider the amount of capital injected in the Italian province by Foreign Direct Investments instead of considering just the number of ICT-related investments. To this end, more than only one probit regression has been made, considering two different independent variables: "fdi_ict" for the number of ICT-related FDIs, and "ict_capital" for the amount of capital injected in the Italian province when the investment is in the field of ICT.

4.2.3 Control variables

Finally, a set of control variables has been included in the analysis. The addition of these control variables is due to the fact that the error term ε_{it} may not include factors that can have an influence (good or bad) on the dependent variable. Indeed, the error arises because of factors, or variables, that influence the dependent variable but are not included in the regression function. In general, there are always "omitted variables" and when it is the case, the final result of the estimation is said to be biased.
The control variables included in the analysis are:

- "patents_pre", that considers the number of ICT-related patents developed in the province, in the first subperiod (1999-2002, before the injection of capital through FDIs);
- "specICTl1", that checks whether the considered province was already specialized in ICT in the subperiod before the injection of capital through FDIs;

- “comple_fdi”, that reports the number of Foreign Direct Investments in industry sectors complementary to ICT that have taken place in that subperiod;
- “comple_capital”, that specifies the amount of capital injected in industry sectors complementary to ICT through FDIs, in that subperiod.

4.3 Descriptive Statistics

The main aspects of the final dataset are now analyzed more specifically. Starting from the categorial variables, the frequencies of each observation in case of subperiods (identified by the variable “subperiod”) and in case of Italian provinces (identified by the variable “sigla”) and are reported in Figure 6 and Figure 7, respectively.

subperiod	Freq.	Percent	Cum.
2003-2006	110	25.23	25.23
2007-2010	110	25.23	50.46
2011-2014	110	25.23	75.69
2015-2019	106	24.31	100.00
Total	436	100.00	

Figure 6: Frequencies of observations in subperiods.

As shown in Figure 6, the number of observations is almost identical for each subperiod considered, with a number of observations equal to 110. Only the last subperiod, from 2015 to 2019 has 4 missing observations due to some missing observations in the original source (that is, the AIDA database). In Figure 7 it is possible to see which are the four missing provinces in the last subperiod, being the frequency of each province equal to 4 and only the frequencies of Carbonia-Iglesias (CI), Ogliastra (OG), Olbia-Tempio (OT) and Medio Campidano (VS) are equal to 3.

sigla	Freq.	Percent	Cum.
AG	4	0.92	0.92
AL	4	0.92	1.83
AN	4	0.92	2.75
AO	4	0.92	3.67
AP	4	0.92	4.59
AQ	4	0.92	5.50
AR	4	0.92	6.42
AT	4	0.92	7.34
AV	4	0.92	8.26
BA	4	0.92	9.17
BG	4	0.92	10.09
BI	4	0.92	11.01
BL	4	0.92	11.93
BN	4	0.92	12.84
BO	4	0.92	13.76
BR	4	0.92	14.68
BS	4	0.92	15.60
BT	4	0.92	16.51
BZ	4	0.92	17.43
CA	4	0.92	18.35
CB	4	0.92	19.27
CE	4	0.92	20.18
CH	4	0.92	21.10
CI	3	0.69	21.79
CL	4	0.92	22.71
CN	4	0.92	23.62
CO	4	0.92	24.54
CR	4	0.92	25.46
CS	4	0.92	26.38
CT	4	0.92	27.29
CZ	4	0.92	28.21
EN	4	0.92	29.13
FC	4	0.92	30.05
FE	4	0.92	30.96
FG	4	0.92	31.88
FI	4	0.92	32.80
FM	4	0.92	33.72
FR	4	0.92	34.63
GE	4	0.92	35.55
GO	4	0.92	36.47
GR	4	0.92	37.39
IM	4	0.92	38.30
IS	4	0.92	39.22
KR	4	0.92	40.14
LC	4	0.92	41.06
LE	4	0.92	41.97
LI	4	0.92	42.89
LO	4	0.92	43.81
LT	4	0.92	44.72
LU	4	0.92	45.64
MB	4	0.92	46.56
MC	4	0.92	47.48
ME	4	0.92	48.39
MI	4	0.92	49.31
MN	4	0.92	50.23
MO	4	0.92	51.15
MS	4	0.92	52.06
MT	4	0.92	52.98
NA	4	0.92	53.90
NO	4	0.92	54.82
NU	4	0.92	55.73
OG	3	0.69	56.42
OR	4	0.92	57.34
OT	3	0.69	58.03
PA	4	0.92	58.94
PC	4	0.92	59.86
PD	4	0.92	60.78
PE	4	0.92	61.70
PG	4	0.92	62.61
PI	4	0.92	63.53
PN	4	0.92	64.45
PO	4	0.92	65.37
PR	4	0.92	66.28
PT	4	0.92	67.20
PU	4	0.92	68.12
PV	4	0.92	69.04
PZ	4	0.92	69.95
RA	4	0.92	70.87
RC	4	0.92	71.79
RE	4	0.92	72.71
RG	4	0.92	73.62
RI	4	0.92	74.54
RM	4	0.92	75.46
RN	4	0.92	76.38
RO	4	0.92	77.29
SA	4	0.92	78.21
SI	4	0.92	79.13
SO	4	0.92	80.05
SP	4	0.92	80.96
SR	4	0.92	81.88
SS	4	0.92	82.80
SV	4	0.92	83.72
TA	4	0.92	84.63
TE	4	0.92	85.55
TN	4	0.92	86.47
TO	4	0.92	87.39
TP	4	0.92	88.30
TR	4	0.92	89.22
TS	4	0.92	90.14
TV	4	0.92	91.06
UD	4	0.92	91.97
VA	4	0.92	92.89
VB	4	0.92	93.81
VC	4	0.92	94.72
VE	4	0.92	95.64
VI	4	0.92	96.56
VR	4	0.92	97.48
VS	3	0.69	98.17
VT	4	0.92	99.08
VV	4	0.92	100.00
Total	436	100.00	

Figure 7: Frequencies of observations in Italian provinces.

Among the continuous variables, the most important ones have been summarized in the following. More specifically, those related to the variables included in the probit regression model have been selected, like the coefficient of localization, the number of patents and the capital invested.

Variable	Obs	Mean	Std. Dev.	Min	Max
ict_capital	**436**	**70.72712**	**589.3158**	**0**	**8359.335**

Figure 8: Variable ict_capital summarized.

The variable *ict_capital*, reporting the amount of capital injected in Italian provinces through ICT-related FDIs, shows a mean value of 70.72 but the standard deviation is very high (589.31). This is due to the fact that there are provinces that do not receive capital (as shown by the minimum value of zero), whereas other provinces (like Milan) attract investments a lot more than others (and indeed, the maximum value is far higher than zero).

Variable	Obs	Mean	Std. Dev.	Min	Max
comple_cap~l	**436**	**28.31259**	**188.9556**	**0**	**2708.1**

Figure 9: Variable comple_capital summarized.

As for the previous case, the standard deviation for the variable *comple_capital* reported in Figure 9 is high as well. Being this variable reporting the amount of capital injected in Italian provinces in complementary sectors to ICT, the reasoning is the same as before (having a minimum value of zero, and a maximum value of 2708.1).

Variable	Obs	Mean	Std. Dev.	Min	Max
patents_pre	**408**	**215.5784**	**387.7262**	**1**	**3101**

Figure 10: Variable patents_pre summarized.

The variable *patents_pre* shows the number of ICT-related patents in the first subperiod (1999-2002) before the first injection of capital through FDIs that took place in the second one (2003-2006). The mean value, as reported in Figure 10, is about 215. Having about 215 ICT-related patents in each Italian province before FDIs would be great for the development of ICT specializations, but as for the previous cases, the standard deviation is very high (about 387). As for the capital invested, there is a substantial heterogeneity among Italian provinces on patents field as well. There are provinces with one single ICT-related patent only, as some others with way more with a maximum value of 3101 (again, the case of Milan).

Variable	Obs	Mean	Std. Dev.	Min	Max
ict_patents	**408**	**222.777**	**359.1266**	**1**	**3282**

Figure 11: Variable ict_patents summarized.

Same situation for the number of ICT-related patents in the following subperiods, despite the injection of capital through Foreign Direct Investments. The same reasoning as before applies, considering for example the huge difference between the minimum and the maximum number of ICT patents in Italian provinces shown in Figure 11.

(mean) spec

	Percentiles	Smallest		
1%	.1783756	0		
5%	.3091396	.0225798		
10%	.4152078	.0882442	Obs	436
25%	.5671465	.1294606	Sum of Wgt.	436
50%	.7370824		Mean	.7497877
		Largest	Std. Dev.	.296872
75%	.8993743	1.800724		
90%	1.062658	1.805399	Variance	.088133
95%	1.336629	1.806065	Skewness	.8852627
99%	1.756098	1.831506	Kurtosis	4.849438

Figure 12: Variable spec summarized.

The coefficient of localization expressed by the variable *spec*, summarized in Figure 12, shows a mean value of about 0.75 and a standard deviation of about 0.3. As for the previous cases, there are provinces showing very low levels of ICT specialization (some provinces have even no specialization at all, with minimum value equal to zero) whereas others show high specialization in ICT field, with values going from 1 to about 1.83. However, despite the mean value is relatively close to 1 (specialization threshold), the distribution of local specialization is not homogeneous, since about 87% of observations do not show ICT specialization in the single subperiod, and only the remaining 13% is said to be specialized. This means that many observations come close to the threshold at most, but they do not reach it unfortunately.

Finally, the (binary) dependent variable *specICT* is summarized in Figure 13. Being binary, the only admissible values are of course 0 and 1, with no missing values among the 436 observations. The figure shows numerically what it has been said about the coefficient of localization: among the 436 total observations, the value 0 (no specialization) appears in 378 cases and so about 87% of the observations fall on the "non-specialized" side, whereas the value 1, that indicates ICT specialization, appears in the remaining 58 cases (13%).

```
specICT

                  type:  numeric (float)

                 range:  [0,1]                        units:  1
         unique values:  2                        missing .:  0/436

            tabulation:  Freq.  Value
                           378  0
                            58  1
```

Figure 13: Variable specICT summarized.

Considering the outcome from summarizing the variables of interest, the first thing that comes out is that there is no homogeneity among the Italian provinces in terms of attractiveness for Foreign Direct Investments. Several provinces do not receive capital at all in terms of investments in ICT field, whereas all the capital is focused on few specific provinces. As shown in Figure 14, the most attractive Italian province is of course Milan, receiving 16 Foreign Direct Investments in ICT during the considered timespan, from 2003 to 2019, with a total amount of capital of almost 22 billion euros. Besides Milan, the only Italian province that attracts more than one FDI throughout the whole period is Rome (4 FDIs with a total amount of capital of about 4.8 billion euros, way less than Milan). Finally, the only Italian provinces left with at least one FDI are Turin, Palermo, Cagliari, Pisa and Bari, but only Turin received an amount of capital higher than one billion euros (about 2.6 billion euros).

sigla	fdi_ict	ict_capital
MI	16	21901.42
RM	4	4824.17
TO	1	2652.79
PA	1	514.92
CA	1	391.6
PI	1	364.78
BA	1	187.35
CZ	0	0
MC	0	0
BT	0	0
OR	0	0
AO	0	0
PN	0	0

Figure 14: Most targeted Italian provinces by FDIs.

Changing viewpoint and considering now the patents field, the situation is a little bit different. The reigning province is still Milan, with 3101 ICT-related patents from 1999 to 2002 and 10120 from 2003 to 2019. Surprisingly, the second province on the list is Turin, whereas Rome (that is the second targeted province by FDIs) is only fifth. This means that, despite having more ICT-related patents than Rome during the subperiod before the first injection of capital, provinces like Bologna and Monza and Brianza did not receive any capital in the field of ICT by Foreign Direct Investments throughout the whole considered timespan. Therefore, it seems not to exist a direct relationship between the already existing number of ICT patents in a province and the attractiveness of that province for ICT-related FDIs.

sigla	patents_pre	ict_patents
MI	3101	10120
TO	1407	5967
BO	1311	4748
MB	1042	2724
RM	871	3486
VA	779	3174
BG	695	2629
MO	637	2645
VI	596	2664
TV	569	2787
PD	522	2379
CO	496	1880
BS	459	2227
PV	430	1134
FI	394	1780
RE	362	1715
GE	339	1609
LC	328	1212
NO	326	1268
VR	323	1551

Figure 15: Number of ICT-related patents in Italian provinces.

Let us consider now the subperiods individually, checking for the number of Foreign Direct Investments that have been done. Figure 16 shows that, from 2003 to 2006, only one province received a FDI in the ICT sector. That province is Milan, receiving an investment of almost 3.7 billion euros.

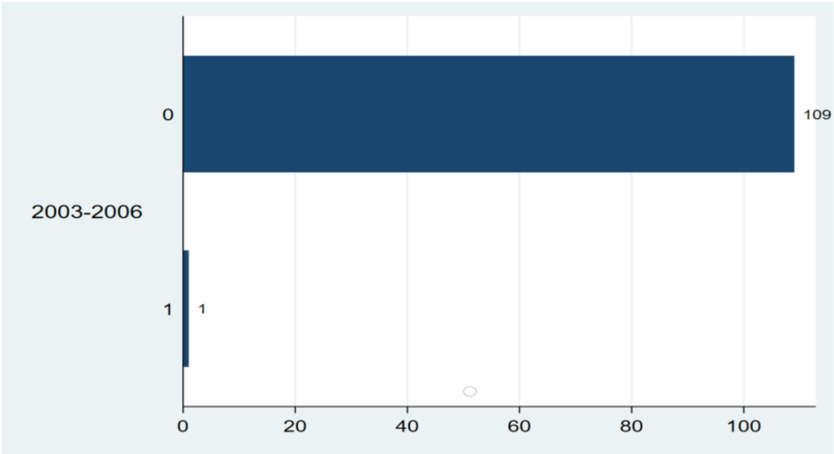

Figure 16: Number of ICT-related FDIs from 2003 to 2006.

Going on from 2007 to 2010, only three Italian provinces receive one single Foreign Direct Investment in ICT. They are Cagliari, Pisa and Turin receiving about 0.39, 0.36 and 2.65 billion euros respectively. Moreover, during this subperiod there have been one province that received four FDIs in ICT. This province is Milan, receiving about 5.8 billion euros.

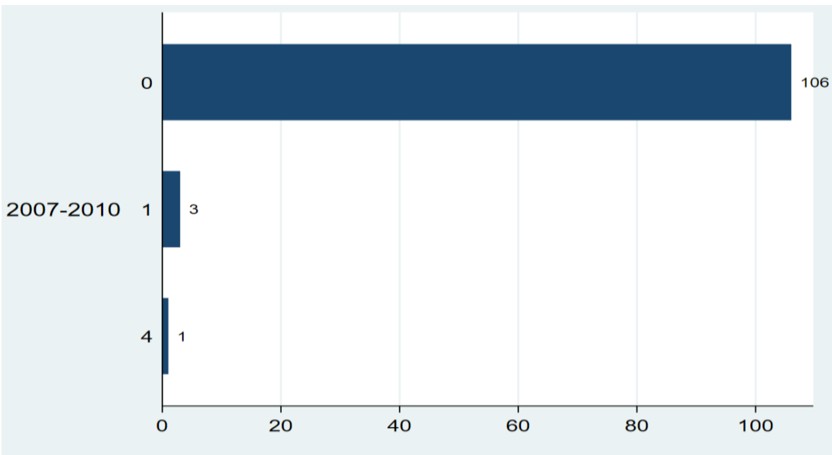

Figure 17: Number of ICT-related FDIs from 2007 to 2010.

During the third subperiod, from 2011 to 2014, three provinces have received FDIs in ICT: Bari received one single FDI of 0.19 billion euros, Milan received four FDIs for a total of 3.9 billion euros and Rome received three FDIs for a total of 2 billion euros.

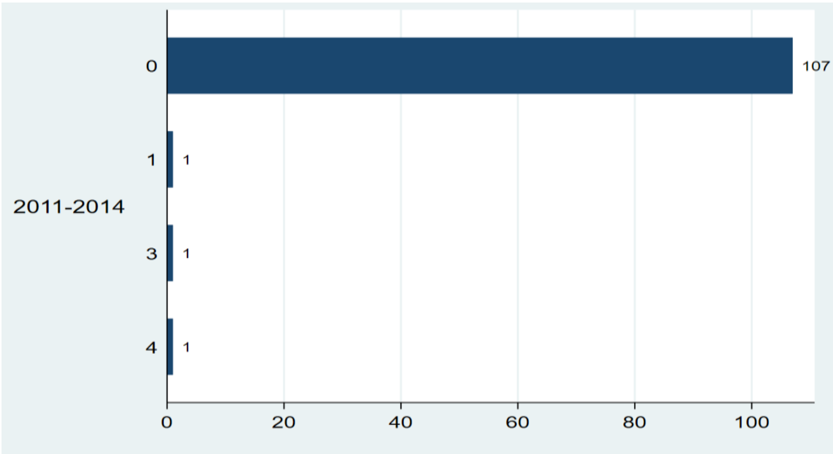

Figure 18: Number of ICT-related FDIs from 2011 to 2014.

During the last subperiod, from 2015 to 2019, there have been three provinces receiving FDIs: Palermo and Rome received only one FDI each for the amount of 0.5 billion euros and 2.7 billion euros respectively, while Milan (of course) received seven FDIs for the total amount of 8.36 billion euros.

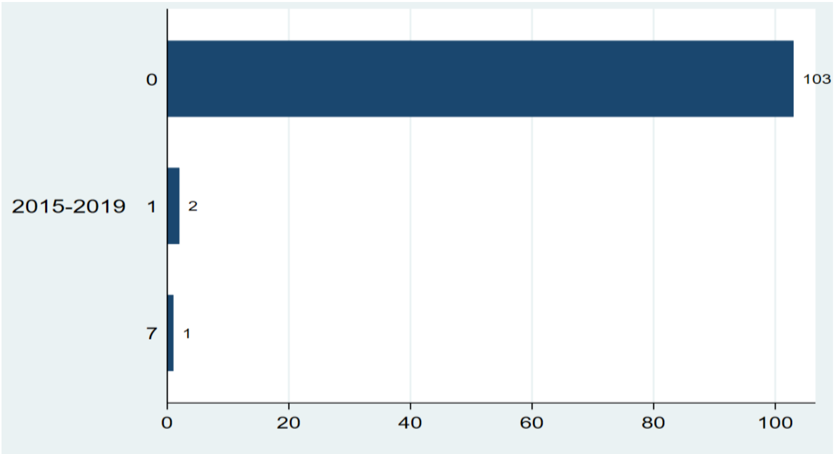

Figure 19: Number of ICT-related FDIs from 2015 to 2019.

Finally, the evolution of ICT specialization in Italian provinces throughout the four considered subperiods, from 2003 to 2019, is shown in Figure 20. On the vertical axis it is possible to notice two subdivisions, both with 0 and 1. The first subdivision refers to the ICT specialization of Italian provinces at time $(t-1)$, identified by the variable *specICTl1*, while the second subdivision refers to the ICT specialization of Italian provinces at time t, identified by the binary dependent variable *specICT*.
As shown in the figure below, the ICT specialization is not very spread among Italian provinces, and it is more likely to be developed (at time t) in provinces that already had an established ICT identity at time $(t-1)$.

Between 2003 and 2006, among the not-specialized provinces in the previous subperiod (1999-2002), only two of them managed to develop an ICT specialization. On the other hand, among those who already exhibited such a specialization, only 4 provinces "lost" their ICT identity whereas 14 kept it.

During the second subperiod (2007-2010), no province has developed an ICT specialization if it was not already ICT-specialized. Only 12 provinces managed to keep their ICT identity that came from the previous subperiod, while 4 provinces even lost it.

The third subperiod (2011-2014) keeps showing the same trend as the previous ones: only 2 provinces managed to get an ICT specialization while not being already specialized in the previous subperiod, whereas 11 provinces kept their ICT identity unchanged.

Lastly, in the fourth subperiod (2015-2019) the maximum number of ICT-specialized provinces has been reached with 17 provinces: 12 of them maintained their ICT identity from the previous subperiod, while 5 of them developed their ICT specialization during the last subperiod considered.

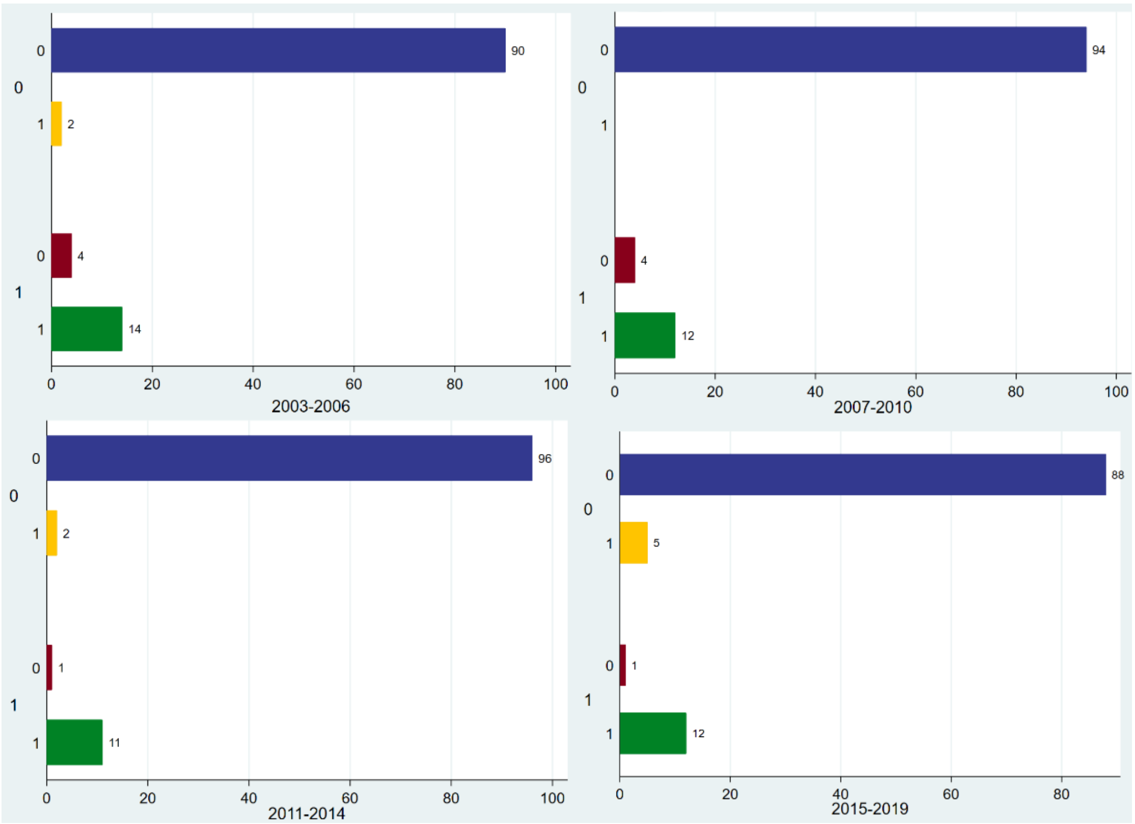

Figure 20: Evolution of ICT specialization in Italian provinces (2003-2019).

5. Results

This chapter will provide an overview of the results obtained from the probit regression model. The analysis has been conducted in two ways: the first one considers the number of Foreign Direct Investments in ICT as independent variable, and the dummy variable *comple_fdi* as control variable for FDIs in complementary sectors to ICT; the second one considers the logarithm of the amount of capital invested in ICT through FDIs as independent variable, and the logarithm of the amount of capital invested in complementary sectors to ICT as one of the control variables.

Let us now take a look at the results.

```
. probit specICT fdi_ict specICT11 pre_spec comple_fdi log_patentspre i.subperiod, robust

Iteration 0:   log pseudolikelihood = -165.00712
Iteration 1:   log pseudolikelihood = -56.500734
Iteration 2:   log pseudolikelihood = -48.179324
Iteration 3:   log pseudolikelihood =  -47.05211
Iteration 4:   log pseudolikelihood = -47.029097
Iteration 5:   log pseudolikelihood = -47.029095

Probit regression                               Number of obs     =        408
                                                Wald chi2(8)      =     106.16
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -47.029095               Pseudo R2         =     0.7150
```

specICT	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
fdi_ict	-.144569	.2503657	-0.58	0.564	-.6352768	.3461389
specICT11	1.906114	.4167968	4.57	0.000	1.089208	2.723021
pre_spec	3.241937	1.069634	3.03	0.002	1.145492	5.338382
comple_fdi	.6499571	.3599248	1.81	0.071	-.0554826	1.355397
log_patentspre	.289707	.1017923	2.85	0.004	.0901978	.4892162
subperiod						
2007-2010	-.6699387	.4170868	-1.61	0.108	-1.487414	.1475364
2011-2014	.0983054	.3900074	0.25	0.801	-.666095	.8627057
2015-2019	.4745002	.3928574	1.21	0.227	-.2954861	1.244487
_cons	-6.194383	1.127968	-5.49	0.000	-8.405161	-3.983606

Figure 21: Results of the probit regression model (first analysis).

As shown by the coefficients in Figure 21, the relationship between the number of ICT-related FDIs in Italy and the ICT specialization is negative. However, this result is not statistically significant since the p-value (0.564) is higher than the significance level (5%). Therefore, it is not possible to reject the H_0 hypothesis that there is no statistically significant relationship between the dependent and the independent variable.

On the other hand, the influence that every other control variable has on the dependent one comes out to be positive. Moreover, the p-value of the control variables is always lower than 5% but for *comple_fdi* (for which it is slightly higher, 7.1%) and so, the obtained results for control variables are statistically significant with a significance level of 5% (significance level of 7.1% for *comple_fdi*). It is then possible to reject the H_0 hypothesis that there is no statistically significant relationship between the dependent variable and the control variables.

About the subperiods, the reference category is the subperiod from 2003 to 2006 and all the results are related to the comparison of the indicated subperiod with the previous one (the subperiod 2007-2010 is compared with the reference category). However, the results related to subperiods are still not statistically significant, reporting a p-value higher than 5% in all the three cases.

Considering instead the second analysis, the results are shown in the following figure.

```
. probit specICT log_kict specICTl1 pre_spec log_kcomple log_patentspre i.subperiod, robust

Iteration 0:   log pseudolikelihood = -165.00712
Iteration 1:   log pseudolikelihood =  -55.83366
Iteration 2:   log pseudolikelihood =   -47.4297
Iteration 3:   log pseudolikelihood = -46.243912
Iteration 4:   log pseudolikelihood = -46.224667
Iteration 5:   log pseudolikelihood = -46.224666

Probit regression                               Number of obs     =        408
                                                Wald chi2(8)      =      84.66
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -46.224666               Pseudo R2         =     0.7199
```

specICT	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
log_kict	-.0756016	.0894693	-0.85	0.398	-.2509582	.099755
specICTl1	1.972795	.4273495	4.62	0.000	1.135205	2.810384
pre_spec	3.224756	1.091301	2.95	0.003	1.085845	5.363667
log_kcomple	.1799148	.0786466	2.29	0.022	.0257703	.3340593
log_patentspre	.3127851	.0984073	3.18	0.001	.1199104	.5056599
subperiod						
2007-2010	-.5807578	.3982314	-1.46	0.145	-1.361277	.1997615
2011-2014	.1860553	.3983422	0.47	0.640	-.5946811	.9667918
2015-2019	.575333	.4061264	1.42	0.157	-.2206602	1.371326
_cons	-6.395552	1.170099	-5.47	0.000	-8.688903	-4.102201

Figure 22: Results of the probit regression model (second analysis).

Comparing the results in Figure 22 with those in Figure 21 it is possible to say that they are pretty similar. Looking at the coefficient of *log_kict*, the relationship between the logarithm of the capital invested in ICT and the ICT specialization in Italian provinces is still negative but not statistically significant, with a *p*-value equal to 0.398. It is then not possible to reject the H_0 hypothesis that there is no statistically significant relationship between the dependent and the independent variable. Moreover, the positive relationship between the dependent and the control variables comes out in the second analysis as well, with the difference that now there is a significance level of 5% for every control variable considered in the analysis, also for the one related to FDIs in complementary sectors, since the p-value is always lower than 5%. Again, it is possible to reject the H_0 hypothesis that there is no statistically significant relationship between the dependent variable and the control variables. The same considerations as before are valid for the subperiods.

A third analysis has been conducted considering the total amount of capital invested in ICT through FDIs and in complementary sectors to ICT (instead of the logarithm of the amount), and the obtained results are essentially identical to the previous two cases.

Overall, the results highlighted a negative non-significant relationship between ICT specialization in Italian provinces and Foreign Direct Investments in ICT, while the specialization in ICT is favored by the number of already existing ICT-related patents in the province, the injection of capital in complementary sectors to ICT and whether the province has an already established ICT identity of not.

5.1 Average Marginal Effects

Among the results shown in Figure 21 and Figure 22, the interpretation of the coefficients can be very difficult. Being the dependent variable binary and so the model a probit regression model, the coefficients can be seen as "log odds". In order to convert them in "odds", it would be useful to calculate the exponential value of the coefficients themselves. For example, the coefficient associated to *log_kcomple* in Figure 22 is 0.18 (rounded up) and so $e^{0.18} \cong 1.20$. This means that the odds that an Italian province develops an ICT specialization following the injection of capital in complementary sectors to ICT through FDIs are 1.20 times higher. This is already more precise than just observing the positive or negative relationship that the mere coefficients report.

Still, one way to convert and interpret the coefficients as percentage points is through the use of the so-called “average marginal effects”. The following figures report the marginal effects associated to both the first and the second analysis.

```
. margins, dydx(*)

Average marginal effects                        Number of obs     =        408
Model VCE     : Robust

Expression    : Pr(specICT), predict()
dy/dx w.r.t. : fdi_ict specICT11 pre_spec comple_fdi log_patentspre 3.subperiod 4.subperiod 5.subperiod
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
fdi_ict	-.0089096	.0152978	-0.58	0.560	-.0388927	.0210736
specICT11	.1174709	.0261616	4.49	0.000	.0661951	.1687467
pre_spec	.1997956	.0653735	3.06	0.002	.0716659	.3279253
comple_fdi	.0400559	.0223465	1.79	0.073	-.0037425	.0838542
log_patentspre	.0178542	.0069484	2.57	0.010	.0042356	.0314728
subperiod						
2007-2010	-.0315437	.0205336	-1.54	0.124	-.0717888	.0087015
2011-2014	.0057706	.0227514	0.25	0.800	-.0388214	.0503626
2015-2019	.032554	.0268938	1.21	0.226	-.0201569	.0852648

Figure 23: Average marginal effects (first analysis).

The interpretation of the coefficients in Figure 23 is the following:

- when one more injection of capital in ICT through FDIs occur in one specific Italian province, the probability of that province to develop an ICT specialization lowers by about 0.89%. However, this result is not statistically significant;
- the fact that an Italian province has already shown an ICT specialization at time $(t-1)$ increases the probability that that province will be specialized in ICT at time t by about 11.75%;
- a one-point higher coefficient of localization will make the probability of the province to develop an ICT specialization increase by about 19.97%;

- when one more injection of capital in complementary sectors to ICT through FDIs occur in one specific Italian province, the probability of that province to develop an ICT specialization increases by about 4%. As said, this result is significant with a significance level of 7.1% (not 5%);
- increasing by one unit the number of ICT-related patents in one Italian province at time $(t-1)$ will increase the probability of that province to develop an ICT specialization by about 1.78%.

```
. margins, dydx(*)

Average marginal effects                        Number of obs     =        408
Model VCE    : Robust

Expression   : Pr(specICT), predict()
dy/dx w.r.t. : log_kict specICTl1 pre_spec log_kcomple log_patentspre 3.subperiod 4.subperiod 5.subperiod
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
log_kict	-.0046052	.0054494	-0.85	0.398	-.0152858	.0060755
specICTl1	.1201699	.0265353	4.53	0.000	.0681617	.172178
pre_spec	.1964312	.0648837	3.03	0.002	.0692614	.323601
log_kcomple	.0109592	.0047939	2.29	0.022	.0015634	.0203551
log_patentspre	.0190528	.0067535	2.82	0.005	.0058163	.0322894
subperiod						
2007-2010	-.026744	.0189577	-1.41	0.158	-.0639004	.0104125
2011-2014	.0105387	.0223391	0.47	0.637	-.033245	.0543225
2015-2019	.0380928	.0267981	1.42	0.155	-.0144305	.0906162

Figure 24: Average marginal effects (second analysis).

Considering instead the second analysis, the interpretation of the coefficients in Figure 24 is similar as before but with some very slight differences:

- when one more injection of capital in ICT through FDIs occur in one specific Italian province, the probability of that province to develop an ICT specialization lowers by about 0.46%. Still, this result is not statistically significant;

- the fact that an Italian province has already shown an ICT specialization at time $(t-1)$ increases the probability that that province will be specialized in ICT at time t by about 12%;
- a one-point higher coefficient of localization will make the probability of the province to develop an ICT specialization increase by about 19.64%;
- when one more injection of capital in complementary sectors to ICT through FDIs occur in one specific Italian province, the probability of that province to develop an ICT specialization increases by about 1.09%. In this case, the result is statistically significant with a significance level of 5%;
- increasing by one unit the number of ICT-related patents in one Italian province at time $(t-1)$ will increase the probability of that province to develop an ICT specialization by about 1.91%.

6. Conclusions

The objective of this thesis paper was to estimate the effect that Foreign Direct Investments in ICT have on the emergence of ICT specializations among Italian provinces. Based on the obtained results, the coefficient of the independent variable related to FDIs in ICT comes out to be negative. This is due to the fact that when MNEs actually invest their capital to bring their activities abroad, this might lead to the so-called "crowding-out" effect. It is true that MNEs bring, through FDIs, valuable technology and knowledge, together with capital, but the entrance in the market of such big entities can drastically lower the piece of the market pie of local businesses, or even force them to leave. For this reason, the process of attracting FDIs has to be carried out carefully when the province has an already established ICT identity, since it might lead to unintended consequences like the exit of local businesses from the industry.
However, the estimations highlighted a non-significant direct effect of FDIs in ICT on the emergence of ICT specializations in Italian provinces. Therefore, it is not possible to reject the H_0 hypothesis that there is no statistically significant relationship between the dependent and the independent variables considered in the analysis.

On the other hand, the main driver of ICT specializations among Italian provinces seems to be the existence of an already established ICT specialization in the province before the injection of capital. Moreover, the number of ICT-related patents in a province before the injection of capital positively relates with the development of a specialization in ICT, as it reflects the actual involvement of that province in the ICT industry sector. Finally, in order to increase (or develop from scratch) the ICT specialization of a province it might be useful to attract Foreign Direct Investments in complementary sectors to ICT, since they would increase the probability of a specialization to emerge by a statistically significant value of about 1.10%.

Further studies will be needed in order to assess more and more precisely the impact that Foreign Direct Investments have on the emergence of new industry specializations. A possible evolution of this analysis might be obtained by adding to the considered model more control variables that would “dry” the results and reduce the error related to omitted variable bias. For example, variables that can be added to ameliorate the model are:

- the relatedness of new technologies to the one that provinces already master, that can be computed as in the paper from Castellani, Marin, Montresor and Zanfei mentioned in the references;
- the economic size and power of each Italian province, computed as a fraction of the Italian GDP;
- the R&D expenses in the ICT sector per Italian province;
- any policies from the Government that favor the development in a specific industry sector (in this case, the ICT sector).

7. References

- (2004). *UN Social Economic Council's Report of the International Telecommunication Union on information and communication technologies statistics*.
- Albornoz, F., Cole, M., Elliott, R., & Ercolani, M. (2009). In search of environmental spillovers.
- Alfaro, L. (2016). Gains from Foreign Direct Investment: Macro and Micro Approaches.
- Alfaro, L., & Charlton, A. (2013). Growth and the Quality of Foreign Direct Investment: Is All FDI Equal?
- Alfaro, L., Chanda, A., Kalemli-Ozcan, S., & Sayek, S. (2004). FDI and Economic Growth: The Role of Local Financial Markets.
- Asongu, S., Akpan, U. S., & Isihak, S. R. (2018). Determinants of foreign direct investment in fast-growing economies: evidence from the BRICS and MINT countries.
- Balland, P. (2016). Relatedness and the geography of innovation.
- Barbieri, N., & Consoli, D. (2019). Regional diversification and green employment in US metropolitan areas.
- Barbieri, N., Perruchas, F., & Consoli, D. (2018). Specialization, diversification and environmental technology-life cycle.
- Belitz, H., & Molders, F. (2016). International knowledge spillovers through high-tech imports and R&D of foreign-owned firms.
- Berge, M. V., & Weterings, A. (2014). Relatedness in eco-technological development in European regions.
- Bilir, K., Chor, D., & Manova, K. (2014). Host-Country Financial Development and Multinational Activity.

- Borensztein, E., Gregorio, J. D., & Lee, J. W. (1998). How Does Foreign Direct Investment Affect Economic Growth?
- Boschma, R. (2015). Proximity and Innovation: a Critical Assessment.
- Bose, T. K. (2012). Advantages and Disadvantages of FDI in China and India.
- Braconier, H., Ekholm, K., & Knarvik, K. M. (2001). In search of FDI-transmitted R&D spillovers: A study based on Swedish data.
- Cainelli, G., Mazzanti, M., & Montresor, S. (2012). Environmental innovations, local networks and internationalization.
- Castellani, D., & Zanfei, A. (2006). Multinational Firms, Innovation and Productivity.
- Castellani, D., Marin, G., Montresor, S., & Zanfei, A. (2020). Foreign Direct Investments and Regional Specialization in Environmental Technologies.
- Colombelli, A., & Quatraro, F. (2019). Green start-ups and local knowledge spillovers from clean and dirty technologies.
- Consoli, D., Castellacci, F., & Santoahla, A. (2019). Do digital skills foster green diversification? A study of European regions.
- Corradini, C. (2019). Location determinants of green technological entry: evidence from European regions.
- Cuervo-Cazurra, A. (2017). ADVANTAGE AND DISADVANTAGE OF FOREIGNNESS AND FOREIGN DIRECT INVESTMENT.
- Dechezlepretre, A., & Glachant, M. (2014). Does Foreign Environmental Policy Influence Domestic Innovation? Evidence from the Wind Industry.
- Elaborazione Centro Documentazione e Studi Anci-Ifel su dati Infocamere. (2012).
- Fu, X. (2008). Foreign direct investment, absorptive capacity and regional innovation capabilities: evidence from China.

- Görg, H., & Greenaway, D. (2004). Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment?
- Komninos, N. (2005). Italian Industrial Districts. *Innovation Clusters*.
- Marchi, V. D., Maria, E. D., Khrishnan, A., & Ponte, S. (2020). Environmental Upgrading in Global Value Chains.
- Marin, A., & Sasidharan, S. (2010). Heterogeneous MNC subsidiaries and technological spillovers: Explaining positive and negative effects in India.
- Markusen, J. R., & Venables, A. J. (1999). Foreign direct investment as a catalyst for industrial development.
- Meivitawanli, B. (2021). Determinants of Foreign Direct Investment: Evidence from Provincial Level Data in Indonesia.
- Montresor, S., & Quatraro, F. (2019). Green technologies and smart specialisation strategies: a European patent-based analysis of the intertwining of technological relatedness and Key-Enabling-Technologies.
- Obalade, T. A. (2014). The Foreign Direct Investment (FDI) Environment in Jordan: A Descriptive Overview.
- Rezza, A. (2013). FDI and pollution havens: Evidence from the Norwegian manufacturing sector.
- Santoalha, A., & Boschma, R. (2019). Diversifying in green technologies in European regions: does political support matter? .
- Stock, J. H., & Watson, M. W. (2014). *Introduction to Econometrics (3rd edition).* Pearson.
- Szanyi, M. (1998). Experiences with Foreign Direct Investments in Eastern Europe - Advantages and Disadvantages.
- Tang, J. (2015). Testing the Pollution Haven Effect: Does the Type of FDI Matter?

- Tanner, A. (2016). The emergence of new technology-based industries: the case of fuel cells and its technological relatedness to regional knowledge bases.
- Todo, Y. (2006). Knowledge spillovers from foreign direct investment in R&D: Evidence from Japanese firm-level data.
- UNCTAD. (2002). World investment report: transnational corporations and export competitiveness. United Nations, New York.
- Xu, B. (2000). Multinational Enterprises, Technology Diffusion, and Host Country Productivity Growth.

8. Appendix

Table 1: Cumulative Standard Normal Distribution.

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\, du$$

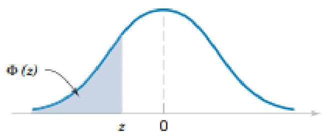

Table I Cumulative Standard Normal Distribution

z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.02	−0.01	−0.00
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687
−3.1	0.000711	0.000736	0.000762	0.000789	0.000816	0.000845	0.000874	0.000904	0.000935	0.000968
−3.0	0.001001	0.001035	0.001070	0.001107	0.001144	0.001183	0.001223	0.001264	0.001306	0.001350
−2.9	0.001395	0.001441	0.001489	0.001538	0.001589	0.001641	0.001695	0.001750	0.001807	0.001866
−2.8	0.001926	0.001988	0.002052	0.002118	0.002186	0.002256	0.002327	0.002401	0.002477	0.002555
−2.7	0.002635	0.002718	0.002803	0.002890	0.002980	0.003072	0.003167	0.003264	0.003364	0.003467
−2.6	0.003573	0.003681	0.003793	0.003907	0.004025	0.004145	0.004269	0.004396	0.004527	0.004661
−2.5	0.004799	0.004940	0.005085	0.005234	0.005386	0.005543	0.005703	0.005868	0.006037	0.006210
−2.4	0.006387	0.006569	0.006756	0.006947	0.007143	0.007344	0.007549	0.007760	0.007976	0.008198
−2.3	0.008424	0.008656	0.008894	0.009137	0.009387	0.009642	0.009903	0.010170	0.010444	0.010724
−2.2	0.011011	0.011304	0.011604	0.011911	0.012224	0.012545	0.012874	0.013209	0.013553	0.013903
−2.1	0.014262	0.014629	0.015003	0.015386	0.015778	0.016177	0.016586	0.017003	0.017429	0.017864
−2.0	0.018309	0.018763	0.019226	0.019699	0.020182	0.020675	0.021178	0.021692	0.022216	0.022750
−1.9	0.023295	0.023852	0.024419	0.024998	0.025588	0.026190	0.026803	0.027429	0.028067	0.028717
−1.8	0.029379	0.030054	0.030742	0.031443	0.032157	0.032884	0.033625	0.034379	0.035148	0.035930
−1.7	0.036727	0.037538	0.038364	0.039204	0.040059	0.040929	0.041815	0.042716	0.043633	0.044565
−1.6	0.045514	0.046479	0.047460	0.048457	0.049471	0.050503	0.051551	0.052616	0.053699	0.054799
−1.5	0.055917	0.057053	0.058208	0.059380	0.060571	0.061780	0.063008	0.064256	0.065522	0.066807
−1.4	0.068112	0.069437	0.070781	0.072145	0.073529	0.074934	0.076359	0.077804	0.079270	0.080757
−1.3	0.082264	0.083793	0.085343	0.086915	0.088508	0.090123	0.091759	0.093418	0.095098	0.096801
−1.2	0.098525	0.100273	0.102042	0.103835	0.105650	0.107488	0.109349	0.111233	0.113140	0.115070
−1.1	0.117023	0.119000	0.121001	0.123024	0.125072	0.127143	0.129238	0.131357	0.133500	0.135666
−1.0	0.137857	0.140071	0.142310	0.144572	0.146859	0.149170	0.151505	0.153864	0.156248	0.158655
−0.9	0.161087	0.163543	0.166023	0.168528	0.171056	0.173609	0.176185	0.178786	0.181411	0.184060
−0.8	0.186733	0.189430	0.192150	0.194894	0.197662	0.200454	0.203269	0.206108	0.208970	0.211855
−0.7	0.214764	0.217695	0.220650	0.223627	0.226627	0.229650	0.232695	0.235762	0.238852	0.241964
−0.6	0.245097	0.248252	0.251429	0.254627	0.257846	0.261086	0.264347	0.267629	0.270931	0.274253
−0.5	0.277595	0.280957	0.284339	0.287740	0.291160	0.294599	0.298056	0.301532	0.305026	0.308538
−0.4	0.312067	0.315614	0.319178	0.322758	0.326355	0.329969	0.333598	0.337243	0.340903	0.344578
−0.3	0.348268	0.351973	0.355691	0.359424	0.363169	0.366928	0.370700	0.374484	0.378281	0.382089
−0.2	0.385908	0.389739	0.393580	0.397432	0.401294	0.405165	0.409046	0.412936	0.416834	0.420740
−0.1	0.424655	0.428576	0.432505	0.436441	0.440382	0.444330	0.448283	0.452242	0.456205	0.460172
0.0	0.464144	0.468119	0.472097	0.476078	0.480061	0.484047	0.488033	0.492022	0.496011	0.500000

Table 2: Cumulative Standard Normal Distribution (continued).

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\, du$$

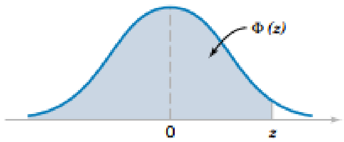

Table II Cumulative Standard Normal Distribution (*continued*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959071	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965621	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967

Table 3: Extract of database AIDA (1999-2019).

	sigla	subperiod	ateco_3d	count_prov_sector	count_prov	count_sector	count_tot	spec	ICT	specICT
1	AG	1999-2002	820	0	6845	8	2439218	0	0	0
2	AG	1999-2002	466	68	6845	38297	2439218	.6327342	0	0
3	AG	1999-2002	691	0	6845	637	2439218	0	0	0
4	AG	1999-2002	252	4	6845	1682	2439218	.8474443	0	0
5	AG	1999-2002	171	0	6845	1384	2439218	0	0	0
6	AG	1999-2002	660	0	6845	112	2439218	0	0	0
7	AG	1999-2002	690	5	6845	3785	2439218	.4707402	0	0
8	AG	1999-2002	221	0	6845	3259	2439218	0	0	0
9	AG	1999-2002	842	0	6845	.	2439218	0	0	0
10	AG	1999-2002	292	13	6845	1770	2439218	2.617262	0	0
11	AG	1999-2002	150	0	6845	86	2439218	0	0	0
12	AG	1999-2002	950	0	6845	.	2439218	0	0	0
13	AG	1999-2002	013	0	6845	183	2439218	0	0	0
14	AG	1999-2002	212	4	6845	1612	2439218	.884244	0	0
15	AG	1999-2002	829	44	6845	21830	2439218	.7182508	0	0
16	AG	1999-2002	107	4	6845	6316	2439218	.225681	0	0
17	AG	1999-2002	420	0	6845	209	2439218	0	0	0
18	AG	1999-2002	324	0	6845	950	2439218	0	0	0
19	AG	1999-2002	562	12	6845	2818	2439218	1.517461	0	0
20	AG	1999-2002	254	0	6845	329	2439218	0	0	0
21	AG	1999-2002	612	0	6845	36	2439218	0	0	0
22	AG	1999-2002	783	0	6845	.	2439218	0	0	0
23	AG	1999-2002	649	15	6845	14421	2439218	.3706577	0	0
24	AG	1999-2002	550	0	6845	86	2439218	0	0	0
25	AG	1999-2002	262	1	6845	4518	2439218	.0788735	0	0
26	AG	1999-2002	650	4	6845	194	2439218	7.347429	0	0
27	AG	1999-2002	749	15	6845	11569	2439218	.4620326	0	0
28	AG	1999-2002	771	3	6845	2147	2439218	.4979278	0	0
29	AG	1999-2002	854	4	6845	1118	2439218	1.274956	0	0
30	AG	1999-2002	590	0	6845	9	2439218	0	0	0
31	AG	1999-2002	370	4	6845	1689	2439218	.8439321	0	0
32	AG	1999-2002	421	44	6845	6880	2439218	2.278985	0	0

Table 4: Extract of database FDI (2003-2019).

	ProjectDate	InvestingCompany	SourceCountry	SourceState	Dest~ry	AdminRegion	IndustrySector	SubSector	Industr~y	Capi~t
1	01mar2011	Karen Millen	United Kingdom	South East …	Italy	Not Specified	Textiles	Clothing &…	Retail	15.5
2	01mar2011	Ashurst	United Kingdom	South East …	Italy	Roma	Business serv…	Legal serv…	Business…	1.4
3	01mar2011	Havaianas	Brazil	Sao Paulo	Italy	Roma	Textiles	Footwear	Retail	15.5
4	01mar2011	YingKe	China	Beijing Mun…	Italy	Verona	Business serv…	Legal serv…	Business…	8.3
5	01mar2011	Yingke Varnai	China	Beijing Mun…	Italy	Not Specified	Business serv…	Legal serv…	Business…	8.3
6	01mar2011	YouTube	United States	California	Italy	Milano	Software & IT…	Internet p…	Sales, M…	4.4
7	01mar2011	Desigual	Spain	Catalonia	Italy	Milano	Textiles	Clothing &…	Retail	15.5
8	01mar2011	Manheim	United States	Georgia	Italy	Milano	Business serv…	Profession…	Business…	8.3
9	01mar2011	Zytech Solar	Spain	Aragon	Italy	Milano	Electronic co…	All other …	Sales, M…	2.3
10	01mar2011	Future Fibres	Spain	Comunidad V…	Italy	Not Specified	Non-automotiv…	Ships & bo…	Sales, M…	2.5
11	01mar2011	Enphase Energy	United States	California	Italy	Not Specified	Electronic co…	All other …	Sales, M…	2.3
12	01mar2011	Axiros	Germany	Bayern	Italy	Milano	Software & IT…	Software p…	Sales, M…	4.4
13	01mar2011	Information Bui…	United States	New York	Italy	Milano	Software & IT…	Software p…	Sales, M…	4.4
14	01mar2011	Hypo Alpe-Adria…	Austria	Sudosterrei…	Italy	Not Specified	Financial ser…	Retail ban…	Business…	30.7
15	01mar2011	Coventor	United States	North Carol…	Italy	Not Specified	Software & IT…	Software p…	Sales, M…	4.4
16	01mar2011	Sonae Sierra	Portugal	Portugal	Italy	Not Specified	Real estate	Commercial…	Construc…	113.1
17	01mar2011	Ashfield Meetin…	Ireland	Ireland	Italy	Roma	Business serv…	Other supp…	Business…	8.3
18	01mar2011	Deufol Group (D…	Germany	Hessen	Italy	Mantova	Paper, printi…	Converted …	Manufact…	37.9
19	01mar2011	Smartclip	Germany	Hamburg	Italy	Milano	Business serv…	Advertisin…	Business…	8.3
20	01mar2011	Buro Happold	United Kingdom	South West …	Italy	Milano	Business serv…	Heavy & ci…	Business…	4.2
21	01feb2011	Chemtura	United States	Pennsylvania	Italy	Latina	Chemicals	Basic chem…	Manufact…	39.1
22	01feb2011	Essar Steel	India	Maharashtra	Italy	Milano	Metals	Steel prod…	Sales, M…	1
23	01feb2011	Rituals	Netherlands	West-Nederl…	Italy	Not Specified	Consumer prod…	Cosmetics,…	Retail	89.1
24	01feb2011	Rituals	Netherlands	West-Nederl…	Italy	Not Specified	Leisure & ent…	Performing…	Construc…	76.2
25	01feb2011	Blancco	Finland	Eastern Fin…	Italy	Not Specified	Software & IT…	Software p…	Sales, M…	4.4
26	01feb2011	Wall Street Ins…	United Kingdom	South East …	Italy	Trento	Business serv…	Profession…	Educatio…	14.6
27	01feb2011	Stradivarius	Spain	Galicia	Italy	Milano	Textiles	Clothing &…	Retail	15.5
28	01feb2011	Tuc Tuc	Spain	La Rioja	Italy	Not Specified	Textiles	Clothing &…	Retail	15.5
29	01feb2011	Kia Motors	South Korea	Seoul	Italy	Not Specified	Automotive OEM	Automobiles	Sales, M…	5.5
30	01feb2011	A Schulman	United States	Ohio	Italy	Varese	Plastics	Urethane, …	Manufact…	23.7
31	01feb2011	Vente-privee.com	France	Ile-de-Fran…	Italy	Torino	Textiles	Clothing &…	Logistic…	23.9
32	01feb2011	Misako	Spain	Catalonia	Italy	Not Specified	Textiles	Apparel ac…	Retail	15.5

Table 5: Extract of the final database.

	sigla	subperiod	ict_patents	patents_pre	CapitalInvestment	JobsCreated	fdi_ict	spec	ICT	specICT	specICTl1	pre_specICT	pre_spec	comple_fdi
1	AG	2003-2006	24	14	219.2	300	0	.2927747	1	0	0	0	.2860372	0
2	AG	2007-2010	16	14	0	0	0	.2747043	1	0	0	0	.2860372	0
3	AG	2011-2014	8	14	0	0	0	.2715461	1	0	0	0	.2860372	0
4	AG	2015-2019	2	14	0	0	0	.2592538	1	0	0	0	.2860372	0
5	AL	2003-2006	292	263	129.3	193	0	.4700938	1	0	0	0	.3997405	0
6	AL	2007-2010	263	263	244.92	293	0	.4926229	1	0	0	0	.3997405	0
7	AL	2011-2014	247	263	214.07	182	0	.5449727	1	0	0	0	.3997405	0
8	AL	2015-2019	134	263	289.976	1126	0	.6771608	1	0	0	0	.3997405	0
9	AN	2003-2006	203	150	76.2	350	0	.9107557	1	0	0	0	.8373037	0
10	AN	2007-2010	347	150	2.3	7	0	.8841169	1	0	0	0	.8373037	0
11	AN	2011-2014	291	150	.5	8	0	.9312913	1	0	0	0	.8373037	0
12	AN	2015-2019	149	150	0	0	0	1.048212	1	1	0	0	.8373037	0
13	AO	2003-2006	36	48	0	0	0	.9781439	1	0	1	1	1.055602	0
14	AO	2007-2010	33	48	63.9	60	0	.9338585	1	0	0	1	1.055602	0
15	AO	2011-2014	38	48	0	0	0	.9450464	1	0	0	1	1.055602	0
16	AO	2015-2019	20	48	16.83	24	0	.8998771	1	0	0	1	1.055602	0
17	AP	2003-2006	44	37	0	0	0	.5586545	1	0	0	0	.5135828	0
18	AP	2007-2010	75	37	430.7	130	0	.4577235	1	0	0	0	.5135828	0
19	AP	2011-2014	72	37	0	0	0	.5250642	1	0	0	0	.5135828	0
20	AP	2015-2019	41	37	0	0	0	.6954242	1	0	0	0	.5135828	0
21	AQ	2003-2006	69	54	2708.1	1158	0	1.098332	1	1	0	0	.9996392	1
22	AQ	2007-2010	61	54	62.1	699	0	1.025897	1	1	1	0	.9996392	1
23	AQ	2011-2014	72	54	112.409	111	0	.964625	1	0	1	0	.9996392	0
24	AQ	2015-2019	44	54	0	0	0	.9126468	1	0	0	0	.9996392	0
25	AR	2003-2006	98	94	0	0	0	.743376	1	0	0	0	.7333193	0
26	AR	2007-2010	120	94	2.3	7	0	.7674608	1	0	0	0	.7333193	0
27	AR	2011-2014	130	94	99.9	200	0	.8540847	1	0	0	0	.7333193	0
28	AR	2015-2019	88	94	0	0	0	.7677376	1	0	0	0	.7333193	0
29	AT	2003-2006	86	78	94.3	192	0	.8342665	1	0	0	0	.7479308	0
30	AT	2007-2010	121	78	71.5	62	0	.8620418	1	0	0	0	.7479308	0
31	AT	2011-2014	71	78	1.1	7	0	.9046223	1	0	0	0	.7479308	0
32	AT	2015-2019	38	78	0	0	0	.7511566	1	0	0	0	.7479308	0