

Article

Foreign Doctorate Students in Europe

Lucio Laureti ¹, Alberto Costantiello ², Marco Maria Matarrese ³, and Angelo Leogrande ^{4,*}

¹ Professor of Economics at LUM University-Giuseppe Degennaro, Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union; laureti@lum.it

² Rica Professor of Economics at LUM University-Giuseppe Degennaro, Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union; costantiello@lum.it

³ Assistant Professor of Economics at LUM University-Giuseppe Degennaro, Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union; matarrese.phdstudent@lum.it.

⁴ Assistant Professor of Economics at LUM University-Giuseppe Degennaro, Strada Statale 100 km 18, 70010 Casamassima BA, Puglia, Italy, European Union; leogrande.culture@lum.it

*Corresponding author: leogrande.culture@lum.it

Abstract: The determinants of the presence of "Foreign Doctorate Students" among 36 European Countries for the period 2010-2019 are analyzed in this article. Panel Data with Fixed Effects, Random Effects, WLS, Pooled OLS, and Dynamic Panel are used to investigate the data. We found that the presence of Foreign Doctorate Students is positively associated to "Attractive Research Systems", "Finance and Support", "Rule of Law", "Sales Impacts", "New Doctorate Graduates", "Basic School Entrepreneurial Education and Training", "Tertiary Education" and negatively associated to "Innovative Sales Share", "Innovation Friendly Environment", "Linkages", "Trademark Applications", "Government Procurement of Advanced Technology Products", "R&D Expenditure Public Sectors". A cluster analysis was then carried out through the application of the unsupervised k-Means algorithm optimized using the Silhouette coefficient with the identification of 5 clusters. Finally, eight different machine learning algorithms were used to predict the value of the "Foreign Doctorate Students" variable. The results show that the best predictor algorithm is the "Tree Ensemble Regression" with a predicted value growing at a rate of 114.03%.

JEL Codes: O30; O31; O32; O33; O36

Keywords: innovation and invention; processes and incentives; management of technological innovation and R&D; diffusion processes; open innovation

1. Introduction

This article analyzes the determinants of "Foreign Doctorate Students" in Europe between 2010 and 2019. The data were analyzed using panel econometric techniques, clustering with k-Means algorithm and machine learning algorithms for prediction. The analysis of the condition of foreign doctoral students is necessary given their role in the development of research systems at the national level. However, as evident from the literature briefly reported below, not all countries are ready to develop economic policies capable of allowing the influx of foreign PhD students.

(Fenyves, et al., 2019) analyze the case of foreign students in Hungary. The authors consider not only the role that doctoral education has for Hungary's economic growth, but also the international prestige that Hungary derives from the presence of foreign doctoral students. Therefore, the analysis tends to identify the best educational strategies to offer to international PhD students the best university courses in the education system of Hungary. (Wang & Byram, 2019) proposes an empirical interview-based study to investigate the effectiveness of training of international doctoral students in Chinese universities. The result of the analysis shows that the ability of foreign PhD students to fit into the Chinese university context depends not only on formal education tools but also on informal social and academic networks that facilitate socialization. The supervision of foreign

doctoral students can therefore take place using a mix of formal educational tools and informal social-community interventions. (Jung, 2019) considers the relationship between local and foreign doctorates in Hong Kong universities. The article tries to verify whether international or local doctoral students have different preferences with respect to the supervision and norms of academic institutions. The authors analyzed approximately 500 questionnaires. The results show that foreign PhD students have more confidence in their skills and have been shown to appreciate the relationship with their supervisors more than local students. In addition, international doctoral students have also expressed greater appreciation for university institutions. This analysis shows that if universities are interested in activating international doctoral courses, they must also consider the possibility of using specific pedagogical approaches for international doctoral students. (Doyle, et al., 2018) analyzes the case of African doctoral students present in New Zealand highlighting the difficulties in writing the thesis and in relations with the supervisors. (Li & Qi, 2019) investigate the motivations that drive Chinese PhDs to go abroad. The study is carried out through 22 interviews. The results show the presence of four different reasons that push doctorates to go abroad, namely: aspects related to the university environment, individual conditions, the need to gain experience abroad and relationships with family and friends. The analysis of the motivations that push Chinese doctorates to go abroad can be useful for higher education institutions in shaping doctoral study programs.

(Lee, et al., 2018) analyze at the case of Chinese PhD students in Australia. The authors point out Chinese doctoral students are highly regarded for their research contributions. Universities try to attract and retain Chinese doctoral students. The study was carried out through interviews conducted in Australian universities. The results show that Chinese PhD students are very sensitive to the social, spiritual, and family dimensions. Therefore, if Australian universities are interested in such doctoral students, they must offer working conditions that meet the social, cultural, and spiritual needs of Chinese doctoral students. (Che Arr, 2021) considers the role of foreign doctoral students in New Zealand. The author analyzes the case of female students from Southeast Asia who moved to New Zealand for their PhDs. The study shows how the relocation led to changes in the lifestyle habits of the respondents. Indeed, living in an urban area has reduced the economic well-being of foreign doctoral students. The author suggests that the New Zealand government introduce new programs for foreign doctorates to improve their living conditions.

(Wang, et al., 2020) consider the role of Chinese PhDs abroad, and particularly in North America in the agricultural economics sector. The analysis shows the very important role those Chinese doctoral students have played in North America in the development of the discipline of agricultural economics thanks to their commitment to scientific research. Chinese doctorates abroad are increasingly gaining access to the role of professor in American universities thanks to their publications in prestigious scientific journals. Finally, many Chinese PhD students who trained overseas have returned to China and are actively participating in the development of China's agrarian economy.

(Mathies, 2021) analyze the case of international students in Finland trying to understand what are the reasons that can push international students to stay in Finland even after completing their studies. The analysis shows that international students have more opportunities to stay in Finland after completing their studies if they have family ties or if they have significant opportunities in the job market. The analysis carried out by the authors does not specifically refer to foreign doctoral students, however the same considerations can also be considered valid for international doctoral students. The South Korean university system has made extensive use of international academic exchanges. Many of South Korea's university professors have taken PhDs from US universities. In addition, foreign doctoral students are gladly accepted in South Korea. (Lim & Shin, 2018) confirm the role of international relations among national academic systems thanks also to the role of foreign doctoral students. (Amiri & Puteh, 2018) tackle the topic of anxiety in oral communication in foreign PhD students in universities in Malaysia. Analysis shows that foreign PhD students show high levels of communication anxiety. The author

suggests that academic institutions pay particular attention to the problem of oral communication apprehension as this problem can have a negative impact on the performance of foreign doctoral students.

(Baloch, et al., 2021) takes into consideration the role of foreign doctorates in Pakistani universities. The analysis takes into consideration the performance differences in terms of research productivity between foreign and Pakistani doctorates. The data used was collected from 14 Pakistani universities. 241 questionnaires were carried out. The results show that the publications of foreign doctorates from a quantitative point of view are not superior to the publications of Pakistani doctorates. However, foreign doctoral publications are qualitatively better than Pakistani doctoral publications. Furthermore, foreign doctorates have a higher number of international co-publications than Pakistani doctorates.

(Interiano & Lim, 2018) consider the degree of integration in American universities of eight different doctoral students from various countries. Respondents developed a sense of conflict and pain in accepting the rules of the US university system. This condition is therefore twofold: on the one hand the foreign doctoral students have acquired professional knowledge from US universities, on the other hand the students themselves show the presence of a value dissonance with respect to the norms of the US university system. (Xu & Zhang, 2019) analyze the case of two Chinese PhD students who did their PhD thesis in English in New Zealand. The authors highlight the fact that writing a doctoral thesis in a foreign language has a significant impact also in the scientific research process and becomes a tool to allow a critical dialogue between various cultures. (Xu & Hu, 2020) refer to the relationship between source language and academic dialogue in the case of Chinese PhD students studying in New Zealand. The authors stress the need to create a more welcoming cultural environment to allow foreign doctoral students to develop writing skills aimed at producing scientific articles more efficiently.

(Lee, 2020) analyzes the case of foreign doctoral students in engineering courses in the United States. The author compared the behavior of national doctoral students with international doctoral students. The research key was centered on the question of the sense of belonging of foreign doctoral students. The analysis shows that the sense of belonging of foreign PhDs can be enhanced through the growth of interaction with academic staff and other students. Universities must therefore develop appropriate educational policies to enable foreign engineering doctorates to develop a sense of belonging in the academic community. (Nguyen, 2018) refers to the increasingly widespread practice among Vietnamese students of doing their PhDs abroad. Training of Vietnamese abroad is seen as an important strategy in the context of the country's modernization process. In particular, the analysis takes into consideration the condition of Vietnamese doctoral students in Australia. The author interviewed 38 Vietnamese PhDs in Australia and acquired information about their experiences, expectations, and their personal and professional future. The results of the analysis show that Vietnamese PhDs who studied in Australia have re-imagined their own experience, their own cultural identity, and have often experienced a sense of alienation and suffering. (Kahn, et al., 2019) analyzes the case of sub-Saharan African PhD students studying in South Africa. The results show that 77% of African doctorates studying in South Africa return to their country of origin, while 7% leave the African continent.

(Yang, et al., 2018) consider the growing role of Chinese STEM PhD students studying abroad. The authors interviewed Chinese PhDs who studied at Australian universities. The motivations that drive Chinese students to take PhDs abroad are the need to enrich their life experiences and improve their education and career prospects. However, students often study abroad also thanks to the presence of international cooperation programs between university institutions. Furthermore, Chinese students are also attracted to foreign doctorates who have important and internationally known supervisors in the scientific context. (Hou & Jam, 2020) afford a confrontation between the experience of Chinese and Cameroon doctoral students studying in the United States. The analysis is also carried out considering the identity implications that can give rise to discrimination for

foreign students in the context of neo-nationalist political forms that can have significant impacts on the university system. The authors conducted interviews and discussions to investigate the degree of integration of Chinese and Cameroonian doctoral students in the US. The results of the survey highlight the role of intercultural dialogue for the integration of foreign doctoral students. Finally, a very important role is recognized in educational equity in university institutions to ensure equal opportunities for foreign doctoral students.

(Pappa, et al., 2020) consider the role of stress international PhDs studying in Finland face. The authors conducted interviews with eleven foreign PhDs in Finland. Stress management by PhD students is a crucial factor in their effective integration into the academic community. The main causes of stress for respondents are: competition in the scientific environment, scarcity of research funding, the need to find career prospects and the lack of a support network. However, some have also defined stress as a motivating factor in scientific research. Respondents tried to overcome stress by giving more meaning to research work and trying to fit more effectively into the academic and social environment. Therefore, stress can be understood as a motivating factor in certain circumstances to increase productivity and the inclusion of foreign doctorates in the academic community. (Mao, 2022) addresses in an autobiographical way the difficulties that a foreign doctoral student faces in the path of the research doctorate starting from the assumption that international doctoral students have greater uncertainties. The author believes that compassion, connection, commitment, and creativity can foster the life and existential condition of international doctorates.

(Todoran & Peterson, 2020) refer to the consequences of the 2017 travel ban in the USA on foreign PhD students both from Muslim countries and from other countries in general. The authors verified that foreign doctoral students suffered travel ban as an oppressive, stressful condition that created a hostile climate. Many doctoral students have changed their programs both for work, such as participation in scientific conferences in the sector, and for personal reasons, such as visits to relatives and family members fearing they will never be able to return to the USA. Finally, foreign doctoral students also expressed concerns about their professional future in the USA. (Lopes & Lourenço, 2019) propose a qualitative-quantitative analysis carried out on international doctoral students in Portugal. The authors wondered if PhD students in the humanities and social sciences had a perception of their identity or if they felt more European or international. The results showed that PhD students who studied abroad show a greater aptitude to define themselves as international or European researchers. This analysis therefore makes it possible to verify the existence of a relationship between the personal experiences of doctorates and their conception of identity with respect to the international or European scientific community. (Nomnian, 2018) addresses the case of Thai PhD students studying in Australia. The author verified the presence of a set of cultural, ethnic, and linguistic elements that characterize the social behaviors of Thai PhD students in Australia. Three different elements were found that define the identity of Thai students, namely: identity as an Asian people, identity as people who do not speak English, identity as an ethnic Thai people. Such identity elements can make it very difficult for Thai PhD students to integrate into the Australian university system.

The article continues as follows: the second paragraph presents the econometric model, the third paragraph contains clustering, the fourth paragraph refers to machine learning techniques for prediction, and the fifth paragraph concludes. Finally, the appendix presents the most relevant statistical results

2. The Econometric Model

The value of the variable referred to as "*Foreign Doctorate Students*" was analyzed below with an econometric model. Data from 36 countries were used in the period between 2010 and 2019. The data were analyzed using the panel data models with fixed effects,

panel data with variable effects, WLS, Pooled OLS and dynamic panel. The estimated econometric model is shown below:

$$\begin{aligned}
 & \text{ForeignDoctorateStudents}_{it} \\
 &= a_{it} + b_1(\text{AttractiveResearchSystems})_{it} \\
 &+ b_2(\text{BasicSchoolEntrepreneurialEducationAndTraining})_{it} \\
 &+ b_3(\text{FinanceAndSupport})_{it} \\
 &+ b_4(\text{GovernmentProcurementOfAdvancedTechnologyProducts})_{it} \\
 &+ b_5(\text{InnovationFriendlyEnvironment})_{it} + b_6(\text{InnovativeSalesShare})_{it} \\
 &+ b_7(\text{Linkages})_{it} + b_8(\text{NewDoctorateGraduates})_{it} \\
 &+ b_9(\text{R\&DExpenditurePublicSector})_{it} + b_{10}(\text{RuleOfLaw})_{it} + b_{11}(\text{SalesImpact})_{it} \\
 &+ b_{12}(\text{TertiaryEducation})_{it} + b_{13}(\text{TrademarkApplications})_{it}
 \end{aligned}$$

Where $i=36$ and $t=[2010; 2019]$.

The variable "Foreign Doctorate Students" is defined by a ratio in which the numerator is the number of PhD students from abroad and the denominator is the total number of PhD students. Obviously, the countries that have a higher value of "Foreign Doctorate Students" are more likely to attract researchers to increase technological innovation and invest in Research and Development.

As is evident from the econometric analysis carried out, the value of the "Foreign Doctorate Students" variable is positively associated with:

- *Attractive Research Systems*: The attractiveness of research systems is a composite variable that measures the ability to compete internationally in scientific research. It is a variable made up of three different components namely "International Scientific Co-Publication", "Most Cited Publications" and "Foreign Doctorate Students". Obviously, there is a positive relationship between the presence of "Foreign Doctorate Students" and the attractiveness value of national research systems. In fact, in those European countries where there are major international relations aimed at research in general, the number of foreign doctorates also tends to increase. However, there is an element of criticism that can be made of the positive relationship between "Attractive Research Systems" and "Foreign Doctorate Students" for the fact that "Foreign Doctorate Students" is an active component of the "Attractive Research Systems" macro-variable. We are therefore in the presence of a phenomenon of partial endogeneity. However, it must be considered that the value of "Foreign Doctorate Students" constitutes only one of the three variables of "Attractive Research Systems". European countries that manage to make their research systems attractive are more likely to increase the presence of foreign doctoral students as well.
- *Finance and Support*: It is a composite variable that measures the ability to provide financial support for technological innovation. This variable consists of two essential elements, namely "Venture Capital Expenditures" and "R&D Expenditure in the Public Sector". This variable therefore has the possibility of measuring the ability of a country to support research and development and technological innovation. As evident in the analysis carried out, it appears that there is a positive relationship between the value of the presence of "Foreign Doctorate Students" and the ability of a country to invest in research and development. The motivation for this positive relationship can be easily investigated. In fact, foreign doctoral students are invited to universities and research centers following the preparation of calls for proposals and scholarships funded by public and private bodies. It follows therefore that the greater the ability of a system to finance research, the greater the international offer of calls for tenders and scholarships for foreign students. Therefore, if a country wants to be more attractive from the point of view of research and acquire more foreign PhD students, it must develop a set of economic policies that allow it to finance scholarships for foreign scholarship students. This investment also makes it possible to have positive repercussions also in industrial terms with the creation of start-ups, new cos and university spin-offs that can increase the degree of technological innovation.
- *Rule of Law*: it is a variable that takes into consideration the value of the trust that market operators have in the capacity of a certain legal system to be effective. Market operators are indeed interested in the ability of a legal system to execute contracts, in the protection of private property rights, in the effectiveness of the police and the courts, and in the low probability of suffering criminal violence. This indicator varies from a range of -2.5

to a value of 2.5. It should be noted that countries that have a higher level of "Rule of Law" are also safer countries for both foreign investors and for foreign PhD students. Specifically, countries that have strong legal systems also have the possibility of offering greater guarantees for foreigners who can therefore choose a certain country also as a destination to carry out doctoral studies. There is therefore a positive relationship between the value of the "Rule of Law" and the value of the "Foreign Doctorate Students". This means that if European countries want to increase the presence of foreign doctorates, they must not only act within the university system but on the contrary, they must also act through the improvement of the legal and judicial system and in the strengthening of the police and security systems.

- **Sales Impacts:** It is a macro variable that takes three elements into consideration: "Medium and high-tech product exports", "Knowledge-intensive services exports" and "Sales of new-to market and new-to-firm product innovators". It is therefore a variable that takes into consideration the degree of sophistication present in the products and services that the companies of a certain country have both with reference to the domestic market and to exports. It is obviously necessary that a country that has a productive and industrial system capable of generating innovative products and services is generally also equipped with a very high system of scientific research and technological innovation. That is, highly innovative companies, which produce innovative products and services, generally have the need to use highly qualified human resources who generally also have PhDs. Moreover, it is often the high-tech companies that together with the universities have the funds for the realization of research doctorates. It follows therefore that the European countries where companies are more innovative generally also have greater resources aimed at attracting foreign doctorates. Hence the presence of foreign doctorates is not only a consequence of the organization of the university system but also a need for those European countries that have highly innovative industries.

| | | const | Attracti ve research systems | Basic- school entrepre neurial educatio n and training | Finance and support | Govern ment procure ment of advance d technolo gy product | Innovati on- friendly environ ment | Innovati ve sales share | Linkage s | New doctorat e graduat es | R&D expendit ure public sector | Rule of law | Sales impacts | Tertiary educatio n | Tradem ark applicat ions |
|----------------|-------------|---------|---------------------------------------|--|---------------------------|--|--|-------------------------------|--------------|------------------------------------|--|----------------|------------------|---------------------------|-----------------------------------|
| | | | A1 | A4 | A17 | A22 | A25 | A26 | A33 | A37 | A47 | A48 | A49 | A53 | A56 |
| Pooled OLS | Coefficient | ★ -0.61 | ★ 1.367 | ★ 0.303 | ★ 0.764 | ★ -1.03 | ★ -0.23 | ★ -0.2 | ★ -0.27 | ★ 0.376 | ★ -1.08 | ★ 0.629 | ★ 0.539 | ★ 0.253 | ★ -0.36 |
| | p-value | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** |
| Dynamic Panel | Coefficient | ★ -3.92 | ★ 1.282 | ★ 0.304 | ★ 0.37 | ★ -1.08 | ★ -0.12 | ★ -0.28 | ★ -0.47 | ★ 0.139 | ★ -0.53 | ★ 0.613 | ★ 0.629 | ★ 0.39 | ★ -0.19 |
| | p-value | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** |
| WLS | Coefficient | ★ 0.525 | ★ 1.552 | ★ 0.393 | ★ 0.394 | ★ -1.03 | ★ -0.25 | ★ -0.14 | ★ -0.32 | ★ 0.297 | ★ -0.93 | ★ 0.474 | ★ 0.417 | ★ 0.352 | ★ -0.31 |
| | p-value | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** |
| Fixed Effects | Coefficient | ★ -0.01 | ★ 1.326 | ★ 0.318 | ★ 0.364 | ★ -1.04 | ★ -0.17 | ★ -0.17 | ★ -0.3 | ★ 0.22 | ★ -0.63 | ★ 0.477 | ★ 0.45 | ★ 0.28 | ★ -0.24 |
| | p-value | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** |
| Random Effects | Coefficient | ★ -0.01 | ★ 1.334 | ★ 0.317 | ★ 0.399 | ★ -1.04 | ★ -0.18 | ★ -0.17 | ★ -0.3 | ★ 0.238 | ★ -0.67 | ★ 0.49 | ★ 0.464 | ★ 0.279 | ★ -0.26 |
| | p-value | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** | *** |

- **New Doctorate Graduates:** It is an indicator that considers the number of new PhD students for every 1000 inhabitants out of the total number of graduates. The value is defined as a percentage. The reference population is aged between 25 and 34 years. There is a positive relationship between the number of new PhD students and the number of foreign PhDs. This positive relationship can be better understood by considering that countries that have a greater number of doctoral programs also have the possibility of welcoming a larger number of foreign doctoral students. Furthermore, in European countries where there are major doctoral programs, doctoral programs aimed exclusively at foreigners may also be explicitly indicated. Obviously, there is also the need to reorganize the university system and ensure that there are PhD courses in English to help foreign students and offer university housing solutions that can meet the housing needs of foreign students. It follows therefore that the European countries that are most interested in attracting foreign students must also be able to offer more doctorates, more study programs in a broad sense and act with international university cooperation.
- **Basic School Entrepreneurial Education and Training:** It is an indicator that varies from 1 to 5 and which measures the level of training that primary and secondary schools offer in the subjects of management and management of SMEs. This indicator has a positive association with the value of "Foreign Doctorate Students". This relationship indicates that in countries that attention to management training has long-term effects that tend to characterize

the entire training system up to university and post-graduate training with doctorates. That is, the countries that aim at knowledge as an asset of industrial and technological development, perhaps through innovation and research and development, cultivate this objective during the entire training course, also managing to attract more foreign students. The growth of the managerial culture in the reference population allows countries to have a more innovation-oriented orientation and to enhance the presence of foreign doctorates by also offering them career opportunities.

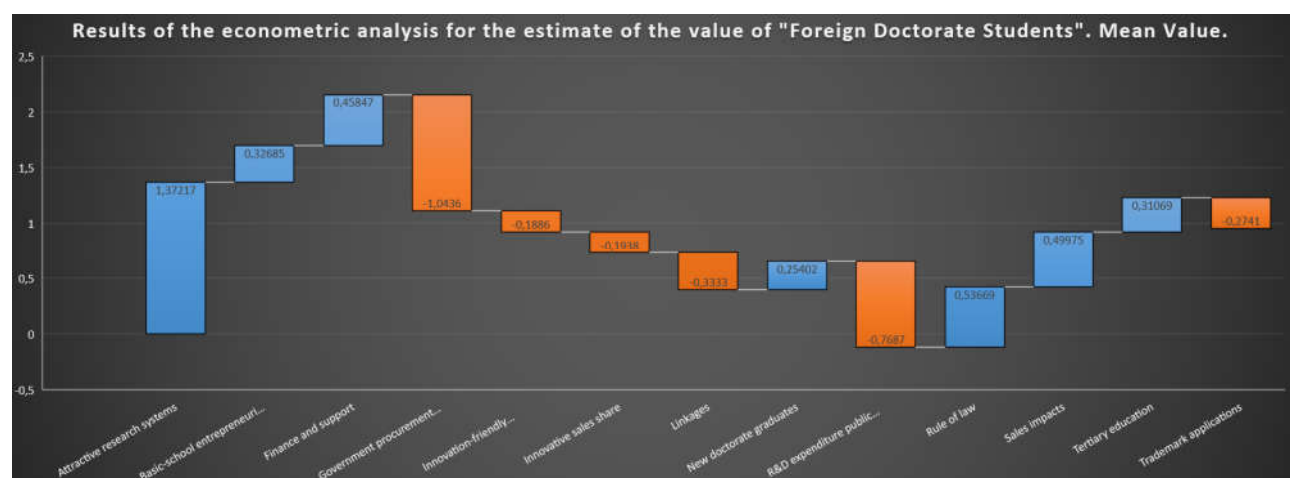
- *Tertiary Education*: represents the percentage of the population aged between 25 and 34 who have completed tertiary education. This indicator is representative of the presence of a part of the population that has advanced professional, technical and scientific skills. However, the indicator does not refer exclusively to training in the technical-scientific field. In fact, notoriously, technological innovation is achieved through the contribution of a wide range of multidisciplinary skills. There is obviously a positive relationship between the percentage of the population having a tertiary education and the number of foreign doctorates. This positive relationship derives from the fact that generally the countries that tend to invest more in training are characterized both by the presence of a large part of the population with tertiary education and by the presence of foreign doctorates. Technological innovation is therefore understood as the product of a human capital widely trained in the professions both in tertiary education and through foreign doctoral students.

The value of Foreign Doctorate Students is instead negatively associated with the following variables:

- *Innovative Sales Share*: It is an indicator that considers the value of new to market and new to business sales as a percentage of turnover. Particularly in the numerator there is the sum of the total turnover deriving from innovative products both for the market and for the company and in the denominator the total turnover for all companies. This indicator can measure the degree of technological innovation that improves the value of the product or even the market in a broad sense. There is a negative relationship between the value of the percentage of sales of innovative products and the value of the presence of "*Foreign Doctorate Students*". This negative relationship may seem in part counterfactual. However, it should be borne in mind that companies may not necessarily be able to produce innovative products through investment in research and development or technological innovation. On the contrary, in very competitive markets, companies can gain competitiveness through process innovations which, however, do not require significant investments in human capital. It follows that if the attention of the firms is to the maximization of the sales, then this activity could also be realized without investment in human capital. This circumstance would justify the negative relationship between sales of innovative products and the presence of foreign doctorates. In fact, it must be considered that to increase sales it is often necessary to reduce market prices. And to reduce prices, process innovations are often preferred to product innovations with negative short-term consequences on human capital.
- *Innovation-Friendly Environment*: It is a macro-variable consisting of two elements: "*Broadband penetration*" and "*Opportunity Driven Entrepreneurship*". It is therefore a variable representative of the presence of an environment that is sensitive to technological innovation in the sense of offering a set of infrastructures and market opportunities to businesses. That is, the indicator does not take into consideration what are the determinants of the human capital employed in the production of technological innovation or of the products created following the investment in Research and Development. For this reason, it appears that countries that invest in the Internet do not necessarily also invest in the training of human capital, especially in the form of doctoral programs reserved for foreigners. In fact, the infrastructural investment in the internet does not necessarily lead to an investment also in the formation of human capital, above all because the dimension of "business innovation" can also be achieved through process innovations that are known to have negative impacts on human capital and therefore also on foreign doctorates.
- *Linkages*: it is a macro-variable made up of three elements, namely "*Innovative SMEs collaborating with others*", "*Public-private co-publications*" and "*Private co-funding of public R&D expenditures*". These variables are particularly capable of maximizing the relationship existing between companies and between the public and private sectors in the sense of technological innovation. In the analysis of these variables, the role of human capital is not considered and consequently neither the role of foreign doctorates. Furthermore, the fact that companies collaborate with each other, or that companies collaborate with the public sector, says nothing about the ability to invest in the formation of human capital and above all in foreign doctorate programs. It follows therefore that

although it is certainly desirable for a country that companies collaborate with the public for technological innovation, this does not necessarily have the capacity to generate investments in human capital and even less in doctoral programs for foreigners.

- *Trademark Applications*: it is a variable that considers the value of trademark filing applications with respect to the value of the Gross Domestic Product. In this ratio, the numerator shows the number of trademark applications filed in Europe and at the World Intellectual Property Office-WIPO, while the denominator is gross domestic product. This indicator is relevant for innovation. In fact, brands are considered as an approximation of the ability to innovate especially in the service sector. There is a negative relationship between the value of trademark registration and the presence of foreign doctorates. This relationship may seem counterfactual. However, it must be considered that although the registration of trademarks can be considered as a highly innovative activity, it is not necessarily a type of innovation that requires a degree of human capital as high as that offered by doctorates. In fact, it should be considered that doctorates are generally oriented towards scientific research, while research and development is not necessarily required to produce trademarks to be registered.
- *Government Procurement of Advanced Technology Products*: It is a variable that public procurement aimed at the acquisition of advanced technological products. The value of this variable varies from a range between 1 and 7. It is therefore a measure that considers the ability of the State and public bodies to express a demand for technological products to be addressed to companies. This variable appears to be negatively associated with foreign doctoral students. This negative relationship may in some way relate to the low quality of innovative products required by the public administration. In fact, the public administration does not have the capacity to acquire complex innovative products containing high levels of research and development. And if the public administration does not purchase innovative products containing outputs generated by research and development, then it will not have the possibility to act through public spending on the determinants of doctoral programs.
- *R&D Expenditure Public Sector*: it represents the value of public investment in research and development. It is a relevant indicator as it highlights the role of public investment in stimulating research and development. However, there is a negative relationship between the value of public investment in research and development and the presence of foreign doctorates. This negative relationship highlights the typical characteristics of public spending on research and development which are obviously not sufficiently oriented towards stimulating foreign relations. In this regard, it is desirable that this relationship will change in the future as public spending on research and development must also relate to a certain degree of internationalization and favor exchange and interaction with foreign doctoral students. Since research doctorates are the most trained professional figures to generate relevant outputs in terms of research and development, it is necessary that European states invest more and more in this sector with particular attention also to international cooperation.



3. Clusterization

A cluster analysis was carried out below using the unsupervised k-Means algorithm. The Silhouette coefficient was used to choose the optimal number of clusters. In particular,

the data of 36 European countries¹ for the period between 2014 and 2021 were used for clustering. The data were acquired from the European Innovation Scoreboard-EIS of the European Commission. To give greater heterogeneity to the cluster analysis, a Silhouette coefficient was selected which corresponds to the following two conditions:

- *maximization of the number of clusters;*
- *Silhouette coefficient value of the single elements of each cluster greater than 0.*

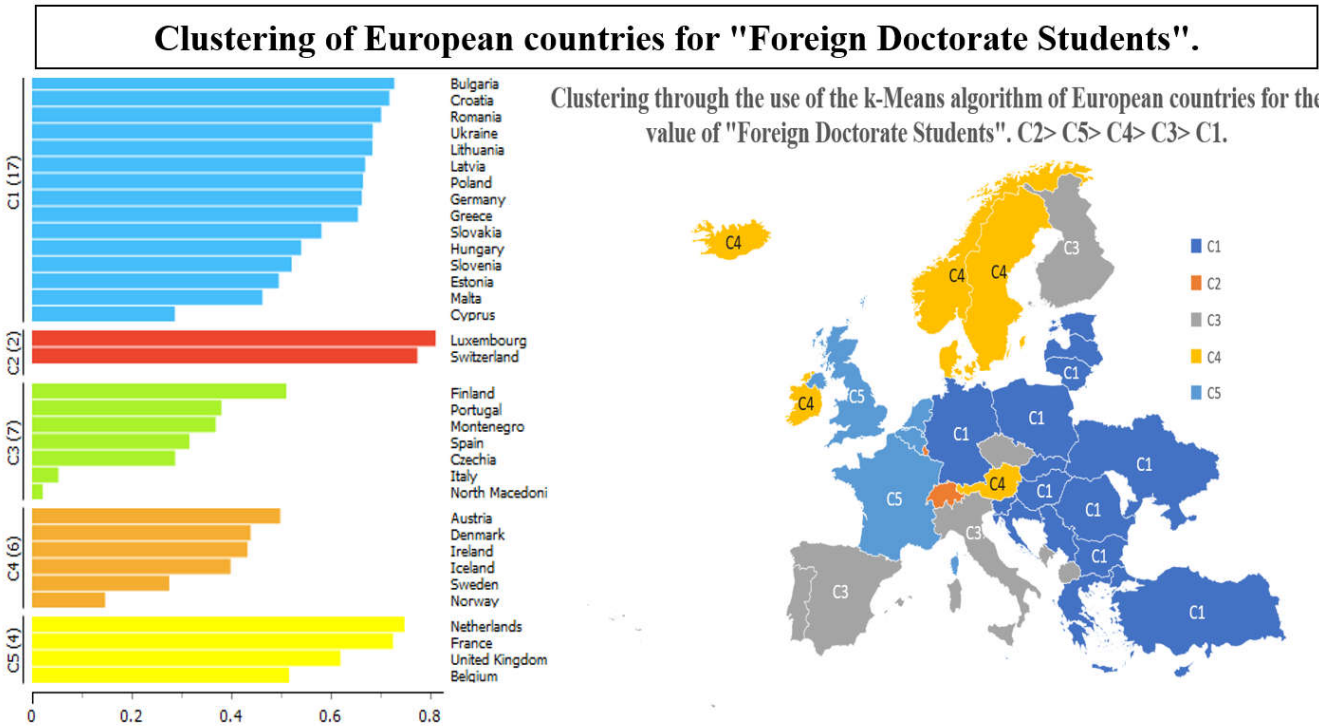
Through the application of this method, it was possible to select five clusters. The five clusters analyzed have the following composition:

- *Cluster 1:* Turkey, Serbia, Bulgaria, Croatia, Romania, Ukraine, Lithuania, Latvia, Poland, Germany, Greece, Slovakia, Hungary, Slovenia, Estonia, Malta, Cyprus;
- *Cluster 2:* Luxembourg, Switzerland;
- *Cluster 3:* Finland, Portugal, Montenegro, Spain, Czechia, Italy, North Macedonia;
- *Cluster 4:* Austria, Denmark, Ireland, Iceland, Sweden, Norway;
- *Cluster 5:* Netherlands, France, United Kingdom, Belgium.

Based on the analysis of the clusters, it is possible to make a ranking based on the median. In particular, cluster 2 has a median value of 353,09, followed by cluster 5 with a value of 264,74, followed by cluster 4 with a median value of 213,01, and by the cluster value 3 with a median value of 116,80 and finally from cluster 1 with a value of 57,90. The following ordering of the clusters was then determined: C2> C5> C4> C3> C1.

As is evident, there is a significant distinction between Northern Europe which has a high level of "Foreign Doctorate Students" and Southern Europe which instead has very low levels of this variable. However, on a more detailed level of analysis, the distinction between Central-Northern Europe and South-Eastern Europe is evident. Specifically, the countries that have the lowest values of this variable in 2021 are Romania with a value equal to 28,45, Poland with an amount equal to 13,47 and Greece with a value equal to 6,87. It follows therefore that at the level of economic policy it is necessary for politicians to intervene to ensure that the countries of Southern and Eastern Europe also have more opportunities to host foreign students. In this regard, for example, it would be appropriate for the European Union to promote degree courses entirely in English to facilitate international exchanges and the integration of Foreign Doctorate Students.

¹ Countries are: Austria, Belgium, Bulgaria, Switzerland, Cyprus, Czechia, Germany, Denmark, Estonia, Greece, Spain, Finland, France, Croatia, Hungary, Ireland, Iceland, Italy, Lithuania, Luxembourg, Latvia, Montenegro, North Macedonia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Sweden, Slovenia, Slovakia, Turkey, Ukraine, United Kingdom.



4. Machine Learning and Predictions

An analysis was carried out below using eight different machine learning algorithms oriented to the prediction of the "Foreign Doctorate Students" variable. The algorithms were trained with 70% of the available data while the remaining 30% was used for prediction. The algorithms were evaluated for their predictive performance through the maximization of R-squared and the minimization of the main statistical errors, namely: "Mean Absolute Error", "Mean Squared Error", "Root Mean Squared Error", "Mean Signed Difference". Following the analysis of the performance of the algorithms, the following order was obtained, that is:

1. *Tree Ensemble Regression* with a payoff value of 5;
2. *Random Forest Regression* with a payoff value of 16;
3. *Linear Regression* with a payoff value of 18;
4. *Gradient Boosted Trees Regression* with a payoff value of 19;
5. *Polynomial Regression* with a payoff value of 21;
6. *PNN-Probabilistic Neural Network* with a payoff value of 33;
7. *Simple Regression Tree and Artificial Neural Network-ANN* with a payoff value of 34.

| Predictions made through the application of the "Tree Ensemble Regression" algorithm | | | | |
|--|-------|------------|--------------------|--------|
| Country | 2021 | Prediction | Absolute Variation | % |
| France | 143,9 | 226,47 | 82,55 | 57,36 |
| Croatia | 133,7 | 106,06 | -27,66 | -20,69 |
| Ireland | 120,8 | 203,43 | 82,65 | 68,43 |
| Luxembourg | 102,1 | 240,91 | 138,80 | 135,93 |
| Montenegro | 93,62 | 142,59 | 48,97 | 52,30 |
| Netherlands | 61,57 | 237,91 | 176,34 | 286,41 |
| Norway | 61,32 | 175,05 | 113,73 | 185,46 |
| Romania | 46,26 | 113,95 | 67,69 | 146,33 |
| Slovenia | 42,4 | 136,27 | 93,87 | 221,38 |
| Slovakia | 28,45 | 140,68 | 112,23 | 394,48 |
| Turkey | 13,47 | 90,85 | 77,38 | 574,48 |
| Mean | 77,06 | 164,92 | 87,87 | 114,03 |

| R-Squared and Statistical Errors with the "Tree Ensemble Regression" algorithm | | | | | |
|--|-----------------|---------------------|--------------------|-------------------------|------------------------|
| | R ² | Mean absolute error | Mean squared error | Root mean squared error | Mean signed difference |
| ANN | -1,847093648242 | 0,409798257817 | 0,248445151272 | 0,498442726170 | 0,253230119789 |
| PNN | -2,434949882417 | 0,508563586748 | 0,324168254568 | 0,569357756220 | 0,029450415441 |
| Simple Regression Tree | -1,834246629188 | 0,514030261908 | 0,318757787067 | 0,564586385833 | 0,166677572619 |
| Gradient Boosted Trees Regression | -0,948281028670 | 0,399014657562 | 0,209026007099 | 0,457193621018 | 0,121672296898 |
| Random Forest | -1,769274238937 | 0,330638995365 | 0,180409280851 | 0,424746136946 | 0,006175983360 |
| Linear Regression | -1,139189453118 | 0,375111459319 | 0,198490590562 | 0,445522828329 | 0,078146488730 |
| Polynomial Regression | -1,168792182323 | 0,358784249948 | 0,190620512103 | 0,436601090359 | 0,203843016353 |
| Tree Ensemble Regression | -0,165172797772 | 0,252762970516 | 0,121616015220 | 0,348734878123 | 0,006169123138 |

Based on the use of the best predictor algorithm or the "Tree Ensemble Regression", the following predictions are made for the countries indicated, namely:

- *France* with an increase from an amount of 143.92 up to a value of 226.47 or an absolute change of 82.55 units equal to 57.36%;
- *Croatia* with a decrease from an amount equal to 133.72 up to a value of 106.06 units equal to a variation of -27.66 units equal to a value of -20.69%;
- *Ireland* with an increase from an amount of 120.78 units up to a value of 203.43 or equal to an absolute variation of 82.65 units equal to a value of 68.43%;
- *Luxembourg* with a change from an amount of 102.11 units up to a value of 240.91 units or equal to a change of 138.80 units equal to 135.93%;
- *Montenegro* with a change from an amount equal to 93.62 units up to a value of 142.59 units equal to a value of 48.97 units equal to 52.30%;
- *Netherlands* with a variation from an amount of 61.57 units up to a value of 237.91 units or equal to a variation of 176.34 units equal to a value of 286.41%;
- *Norway* with an increase from a value of 61.32 units up to a value of 175.05 units or equal to a variation of 113.73 units equal to 185.46%;
- *Romania* with an increase from an amount equal to 46.26 up to a value of 113.95 units or equal to a value of 67.69 or equal to an amount of 146.33%;
- *Slovenia* with a variation from an amount equal to 42.4 units up to a value of 136.27 units or equal to an amount of 93.87 units equivalent to 221.38%;
- *Slovakia* with an increase from an amount of 28.45 units up to a value of 140.68 units or equal to an absolute change of 112.23 units equal to 394.48%;
- *Turkey* with an increase from an amount of 13.47 units up to a value of 90.85 units or equal to a value of 77.38 units equal to 574.48%.

Finally, the average value of the prediction calculated with the "Tree Ensemble Regression" algorithm is expected to grow by 87 units equal to 114.03%. As is evident from the analysis carried out, the value of the variable indicated as "Foreign Doctorate Students" is expected to grow in almost all the countries indicated except in the case of Croatia. The substantially automatic variation predicted by the algorithm is a reverberation of the historical series underlying the prediction which is substantially oriented to the growth of the value of "Foreign Doctorate Students". In fact, in the passage between 2014 and 2021 by

analyzing the historical series, it appears that the value of the presence of "Foreign Doctorate Students" in the 36 European countries has grown by a value of 27.57 units equal to 25.27%. The prediction therefore confirms and strengthens this orientation towards the growth of the value of the analyzed variable, underlining the existence of a significant change within the European nations with reference to the reception of "Foreign Doctorate Students".

5. Conclusions

The determinants of the presence of "Foreign Doctorate Students" among 36 European Countries for the period 2010-2019 are analyzed in this article. Panel Data with Fixed Effects, Random Effects, WLS, Pooled OLS, and Dynamic Panel are used to investigate the data. In the first paragraph, the analysis of the literature highlights the critical issues that university systems must face to effectively welcome international doctoral students. In particular, the issue of performance must be considered. International PhD students are often better and publish better than local PhDs. However, foreign doctorates can also suffer greatly from some identity elements that can be experienced as forms of discrimination. For example, in the case of 2017 travel ban in the USA, many PhD students feared for their condition and professional career. However, the growing attention towards international doctoral students is changing university policies especially in Anglo-Saxon countries and in China. In fact, China tends to present itself internationally as a university system open to foreign doctoral students. The interest in foreign international students is mainly due to the productive capacity and the contribution that these research experts can make in the university and research systems of the host countries.

In the second paragraph an econometric analysis was carried out. We found that the presence of Foreign Doctorate Students is positively associated to "Attractive Research Systems", "Finance and Support", "Rule of Law", "Sales Impacts", "New Doctorate Graduates", "Basic School Entrepreneurial Education and Training", "Tertiary Education" and negatively associated to "Innovative Sales Share", "Innovation Friendly Environment", "Linkages", "Trademark Applications", "Government Procurement of Advanced Technology Products", "R&D Expenditure Public Sectors".

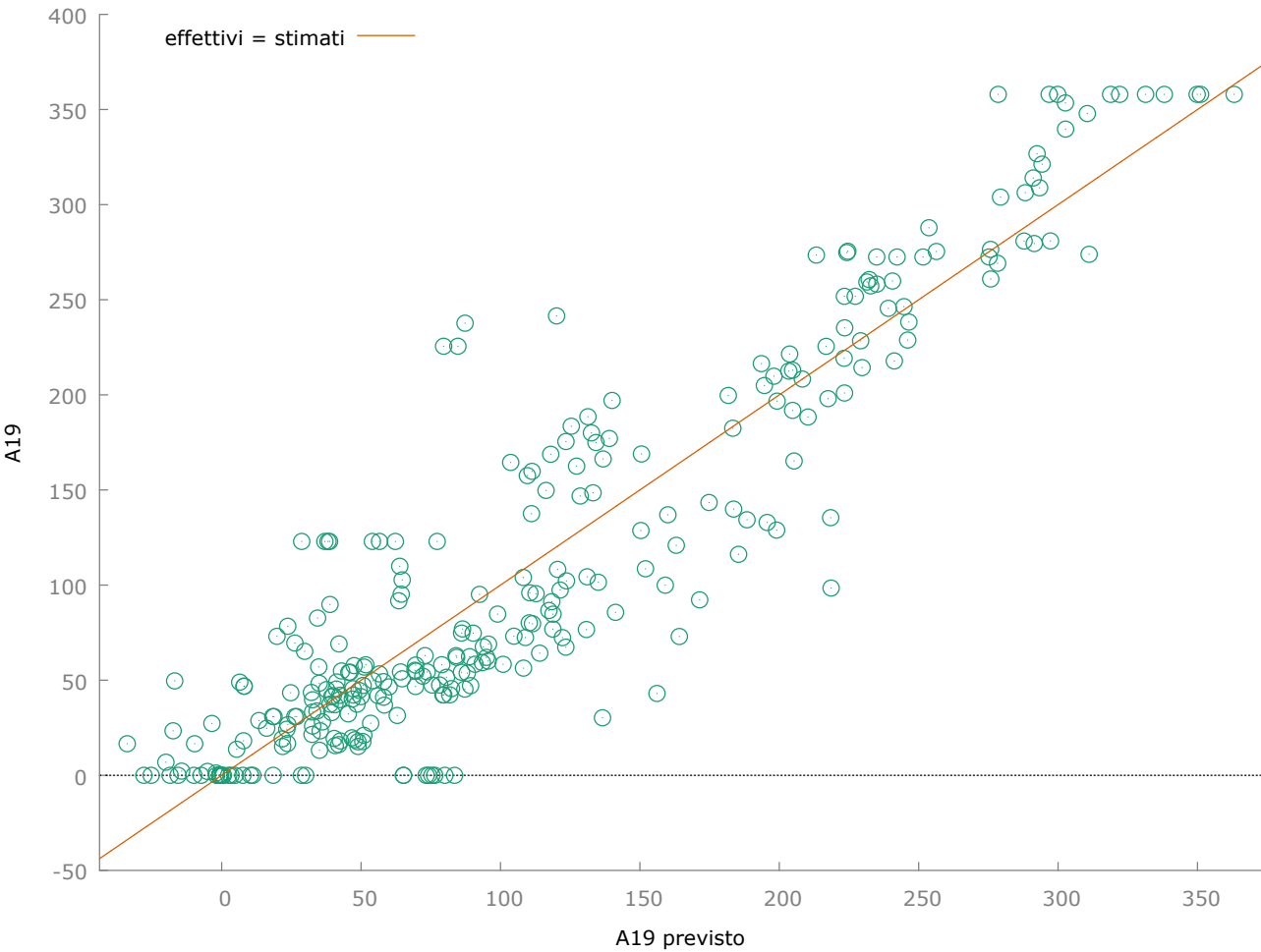
In the third paragraph a cluster analysis was then carried out through the application of the unsupervised k-Means algorithm optimized using the Silhouette coefficient with the identification of 5 clusters. The cluster analysis shows the presence of a clear contrast between the countries of central-northern Europe and the countries of South-Eastern Europe. Finally in the fourth paragraph, eight different machine learning algorithms were used to predict the value of the "Foreign Doctorate Students" variable. The results show that the best predictor algorithm is the "Tree Ensemble Regression" with a predicted value growing at a rate of 114.03%.

6. Appendix

Modello 454: Pooled OLS, usando 360 osservazioni
Include 36 unità cross section
Lunghezza serie storiche = 10
Variabile dipendente: A19

| | Coefficiente | Errore Std. | rapporto t | p-value | |
|-------|--------------|-------------|------------|---------|-----|
| const | -0,607293 | 3,74381 | -0,1622 | 0,8712 | |
| A1 | 1,36661 | 0,0565295 | 24,18 | <0,0001 | *** |
| A4 | 0,302952 | 0,0821246 | 3,689 | 0,0003 | *** |
| A17 | 0,764422 | 0,103845 | 7,361 | <0,0001 | *** |
| A22 | -1,02654 | 0,104524 | -9,821 | <0,0001 | *** |

| | | | | | |
|-----------------------|-----------|------------------------|----------|---------|-----|
| A25 | -0,225779 | 0,0478663 | -4,717 | <0,0001 | *** |
| A26 | -0,203280 | 0,0647614 | -3,139 | 0,0018 | *** |
| A33 | -0,268687 | 0,106917 | -2,513 | 0,0124 | ** |
| A37 | 0,376285 | 0,0581512 | 6,471 | <0,0001 | *** |
| A47 | -1,07589 | 0,0972089 | -11,07 | <0,0001 | *** |
| A48 | 0,629075 | 0,129690 | 4,851 | <0,0001 | *** |
| A49 | 0,539334 | 0,0947776 | 5,691 | <0,0001 | *** |
| A53 | 0,252712 | 0,0413802 | 6,107 | <0,0001 | *** |
| A56 | -0,362119 | 0,0426363 | -8,493 | <0,0001 | *** |
| | | | | | |
| Media var. dipendente | 89,82317 | SQM var. dipendente | 101,2438 | | |
| Somma quadr. residui | 453606,7 | E.S. della regressione | 36,20776 | | |
| R-quadro | 0,876733 | R-quadro corretto | 0,872101 | | |
| F(13, 346) | 189,3004 | P-value(F) | 2,0e-148 | | |
| Log-verosimiglianza | -1795,817 | Criterio di Akaike | 3619,633 | | |
| Criterio di Schwarz | 3674,039 | Hannan-Quinn | 3641,266 | | |
| rho | 0,888549 | Durbin-Watson | 0,307967 | | |



Modello 455: a un passo, usando 288 osservazioni
Incluse 36 unità cross section

Matrice H conforme ad Ox/DPD

Variabile dipendente: A19

| | <i>Coefficiente</i> | <i>Errore Std.</i> | <i>z</i> | <i>p-value</i> | |
|---------|---------------------|--------------------|----------|----------------|-----|
| A19(-1) | 0,124868 | 0,0601382 | 2,076 | 0,0379 | ** |
| const | -3,91676 | 1,71403 | -2,285 | 0,0223 | ** |
| A1 | 1,28177 | 0,183856 | 6,972 | <0,0001 | *** |
| A4 | 0,303771 | 0,0940164 | 3,231 | 0,0012 | *** |
| A17 | 0,370492 | 0,165539 | 2,238 | 0,0252 | ** |
| A22 | -1,07912 | 0,110233 | -9,789 | <0,0001 | *** |
| A25 | -0,123627 | 0,0664022 | -1,862 | 0,0626 | * |
| A26 | -0,278040 | 0,143301 | -1,940 | 0,0523 | * |
| A33 | -0,474184 | 0,141956 | -3,340 | 0,0008 | *** |
| A37 | 0,139171 | 0,0832491 | 1,672 | 0,0946 | * |
| A47 | -0,533308 | 0,136256 | -3,914 | <0,0001 | *** |
| A48 | 0,613254 | 0,159648 | 3,841 | 0,0001 | *** |
| A49 | 0,628911 | 0,263344 | 2,388 | 0,0169 | ** |
| A53 | 0,390115 | 0,129697 | 3,008 | 0,0026 | *** |
| A56 | -0,191195 | 0,0925287 | -2,066 | 0,0388 | ** |

Somma quadr. residui 122168,5 E.S. della regressione 21,15428

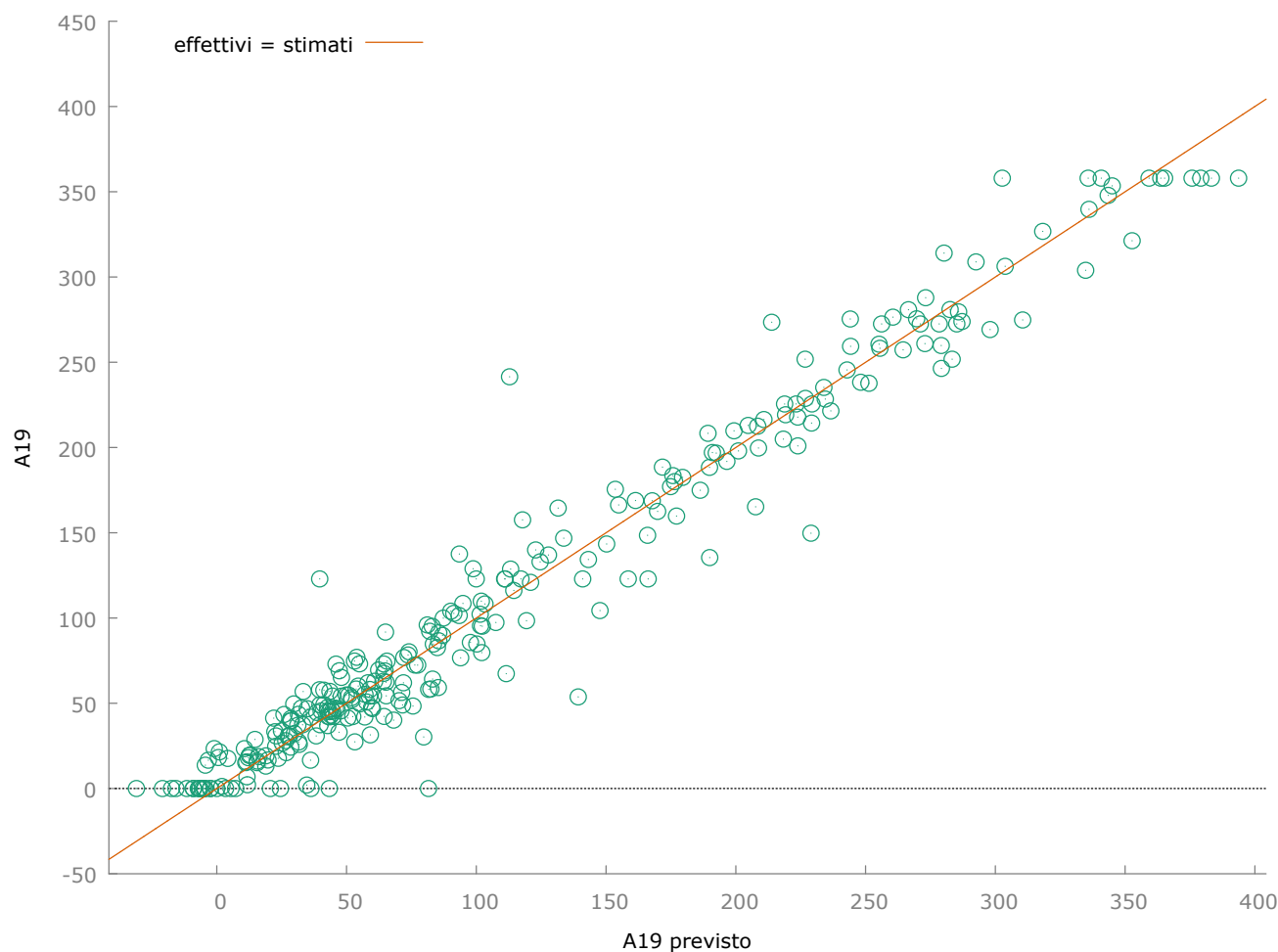
Numero di strumenti = 35

Test per errori AR(1): $z = -1,80629$ [0,0709]

Test per errori AR(2): $z = 0,434239$ [0,6641]

Test di sovra-identificazione di Sargan: Chi-quadro(20) = 46,6436 [0,0007]

Test (congiunto) di Wald: Chi-quadro(14) = 963,744 [0,0000]



Modello 456: WLS, usando 360 osservazioni

Incluse 36 unità cross section

Variabile dipendente: A19

Pesi basati sulle varianze degli errori per unità

| | Coefficiente | Errore Std. | rapporto t | p-value | |
|-------|--------------|-------------|------------|---------|-----|
| const | 0,524900 | 1,58781 | 0,3306 | 0,7412 | |
| A1 | 1,55201 | 0,0341963 | 45,39 | <0,0001 | *** |
| A4 | 0,392941 | 0,0274767 | 14,30 | <0,0001 | *** |
| A17 | 0,394380 | 0,0612143 | 6,443 | <0,0001 | *** |
| A22 | -1,02822 | 0,0558077 | -18,42 | <0,0001 | *** |
| A25 | -0,248981 | 0,0287827 | -8,650 | <0,0001 | *** |
| A26 | -0,142355 | 0,0316942 | -4,492 | <0,0001 | *** |
| A33 | -0,317770 | 0,0635115 | -5,003 | <0,0001 | *** |
| A37 | 0,296853 | 0,0299709 | 9,905 | <0,0001 | *** |
| A47 | -0,933629 | 0,0425069 | -21,96 | <0,0001 | *** |
| A48 | 0,473800 | 0,0615248 | 7,701 | <0,0001 | *** |
| A49 | 0,417132 | 0,0510860 | 8,165 | <0,0001 | *** |
| A53 | 0,351604 | 0,0226702 | 15,51 | <0,0001 | *** |
| A56 | -0,313722 | 0,0255227 | -12,29 | <0,0001 | *** |

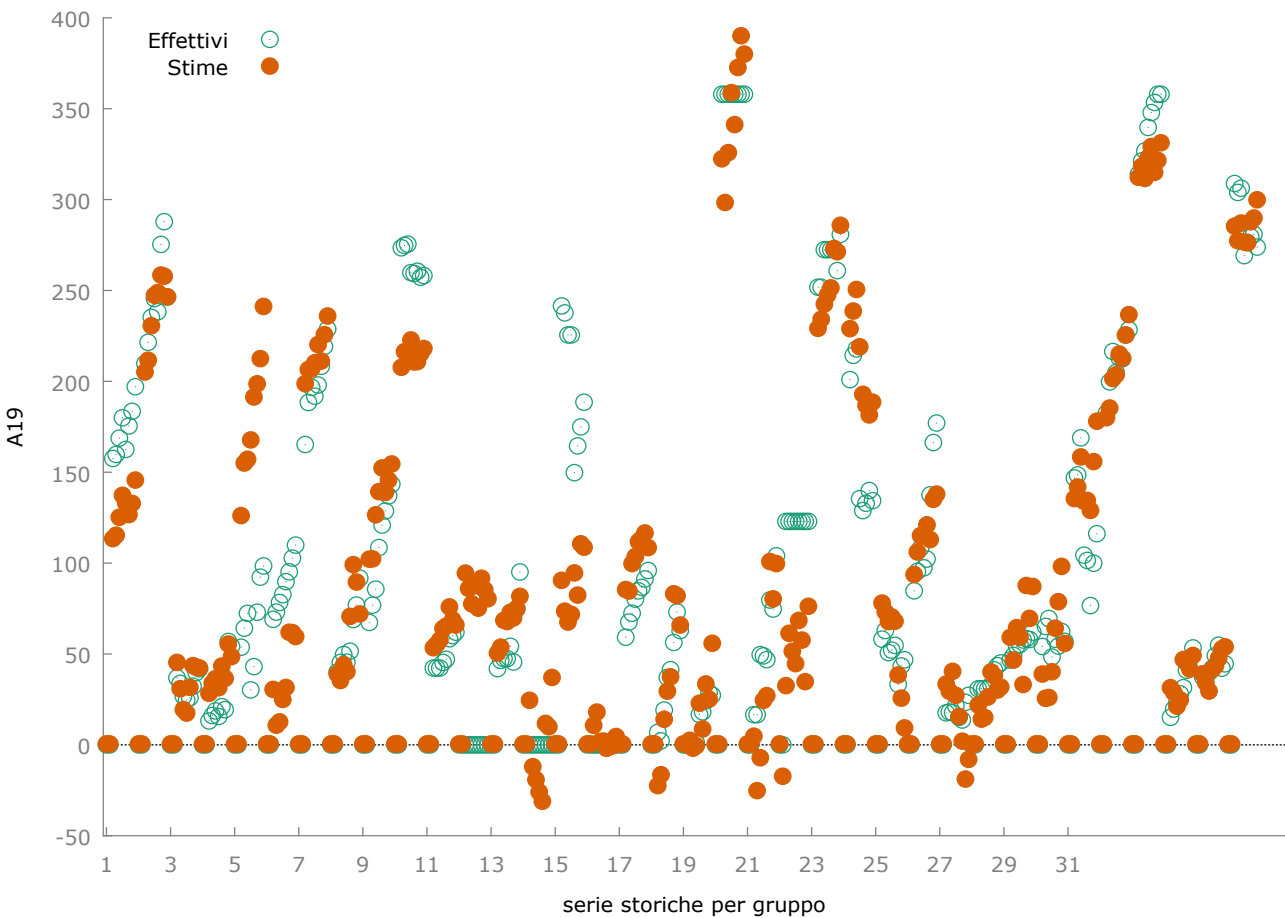
Statistiche basate sui dati ponderati:

| | | | |
|----------------------|-----------|------------------------|----------|
| Somma quadr. residui | 303,8120 | E.S. della regressione | 0,937054 |
| R-quadro | 0,967174 | R-quadro corretto | 0,965940 |
| F(13, 346) | 784,1763 | P-value(F) | 1,3e-247 |
| Log-verosimiglianza | -480,2728 | Criterio di Akaike | 988,5456 |
| Criterio di Schwarz | 1042,951 | Hannan-Quinn | 1010,178 |

Statistiche basate sui dati originali:

| | | | |
|-----------------------|----------|------------------------|----------|
| Media var. dipendente | 89,82317 | SQM var. dipendente | 101,2438 |
| Somma quadr. residui | 497609,6 | E.S. della regressione | 37,92332 |

A19: valori effettivi e stimati



Modello 457: Effetti fissi, usando 360 osservazioni

Incluse 36 unità cross section

Lunghezza serie storiche = 10

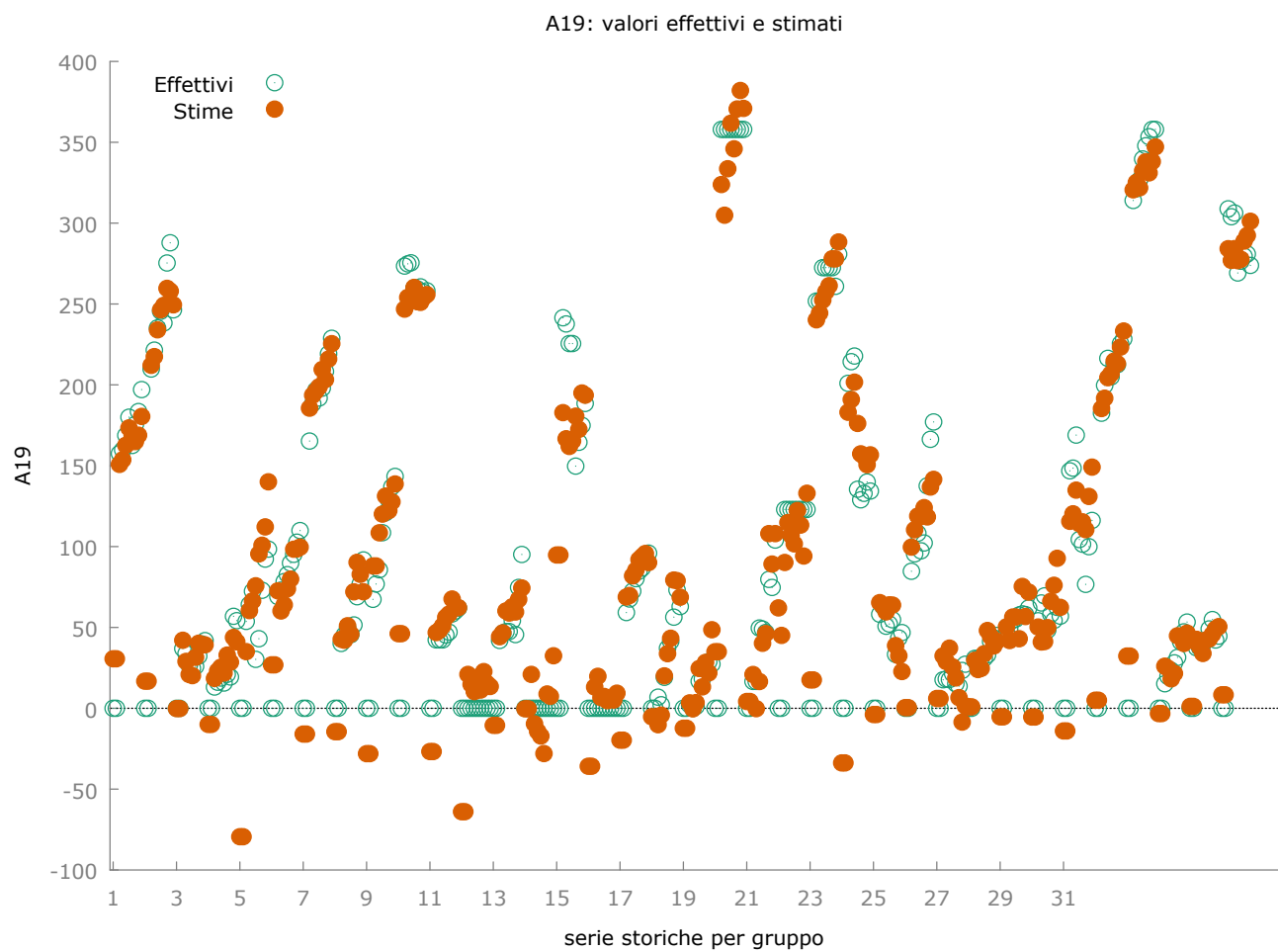
Variabile dipendente: A19

| | Coefficiente | Errore Std. | rapporto t | p-value | |
|-------|--------------|-------------|------------|---------|-----|
| const | -0,0110408 | 2,49228 | -0,004430 | 0,9965 | |
| A1 | 1,32639 | 0,0670620 | 19,78 | <0,0001 | *** |
| A4 | 0,317522 | 0,102862 | 3,087 | 0,0022 | *** |
| A17 | 0,364141 | 0,0923095 | 3,945 | <0,0001 | *** |

| | | | | | |
|-----------------------|-----------|------------------------|----------|---------|-----|
| A22 | -1,04410 | 0,128672 | -8,114 | <0,0001 | *** |
| A25 | -0,168832 | 0,0436602 | -3,867 | 0,0001 | *** |
| A26 | -0,171329 | 0,0623459 | -2,748 | 0,0063 | *** |
| A33 | -0,304176 | 0,106255 | -2,863 | 0,0045 | *** |
| A37 | 0,220169 | 0,0611303 | 3,602 | 0,0004 | *** |
| A47 | -0,626668 | 0,0974780 | -6,429 | <0,0001 | *** |
| A48 | 0,477480 | 0,155554 | 3,070 | 0,0023 | *** |
| A49 | 0,449746 | 0,114971 | 3,912 | 0,0001 | *** |
| A53 | 0,279556 | 0,0468314 | 5,969 | <0,0001 | *** |
| A56 | -0,244071 | 0,0494822 | -4,932 | <0,0001 | *** |
| | | | | | |
| Media var. dipendente | 89,82317 | SQM var. dipendente | 101,2438 | | |
| Somma quadr. residui | 153357,1 | E.S. della regressione | 22,20607 | | |
| R-quadro LSDV | 0,958325 | R-quadro intra-gruppi | 0,889347 | | |
| LSDV F(48, 311) | 148,9908 | P-value(F) | 5,0e-187 | | |
| Log-verosimiglianza | -1600,614 | Criterio di Akaike | 3299,227 | | |
| Criterio di Schwarz | 3489,646 | Hannan-Quinn | 3374,941 | | |
| rho | 0,486969 | Durbin-Watson | 0,788003 | | |

Test congiunto sui regressori -
Statistica test: $F(13, 311) = 192,277$
con $p\text{-value} = P(F(13, 311) > 192,277) = 5,03992e-140$

Test per la differenza delle intercette di gruppo -
Ipotesi nulla: i gruppi hanno un'intercetta comune
Statistica test: $F(35, 311) = 17,3969$
con $p\text{-value} = P(F(35, 311) > 17,3969) = 2,82273e-054$



Modello 458: Effetti casuali (GLS), usando 360 osservazioni

Incluse 36 unità cross section

Lunghezza serie storiche = 10

Variabile dipendente: A19

| | Coefficiente | Errore Std. | z | p-value | |
|-------|--------------|-------------|------------|---------|-----|
| const | -0,00575170 | 6,14432 | -0,0009361 | 0,9993 | |
| A1 | 1,33409 | 0,0627188 | 21,27 | <0,0001 | *** |
| A4 | 0,317058 | 0,0956603 | 3,314 | 0,0009 | *** |
| A17 | 0,398932 | 0,0894453 | 4,460 | <0,0001 | *** |
| A22 | -1,03993 | 0,119755 | -8,684 | <0,0001 | *** |
| A25 | -0,175925 | 0,0420901 | -4,180 | <0,0001 | *** |
| A26 | -0,174173 | 0,0596791 | -2,919 | 0,0035 | *** |
| A33 | -0,301775 | 0,101620 | -2,970 | 0,0030 | *** |
| A37 | 0,237640 | 0,0581605 | 4,086 | <0,0001 | *** |
| A47 | -0,674021 | 0,0933569 | -7,220 | <0,0001 | *** |
| A48 | 0,489818 | 0,145493 | 3,367 | 0,0008 | *** |
| A49 | 0,463627 | 0,107396 | 4,317 | <0,0001 | *** |
| A53 | 0,279467 | 0,0440783 | 6,340 | <0,0001 | *** |

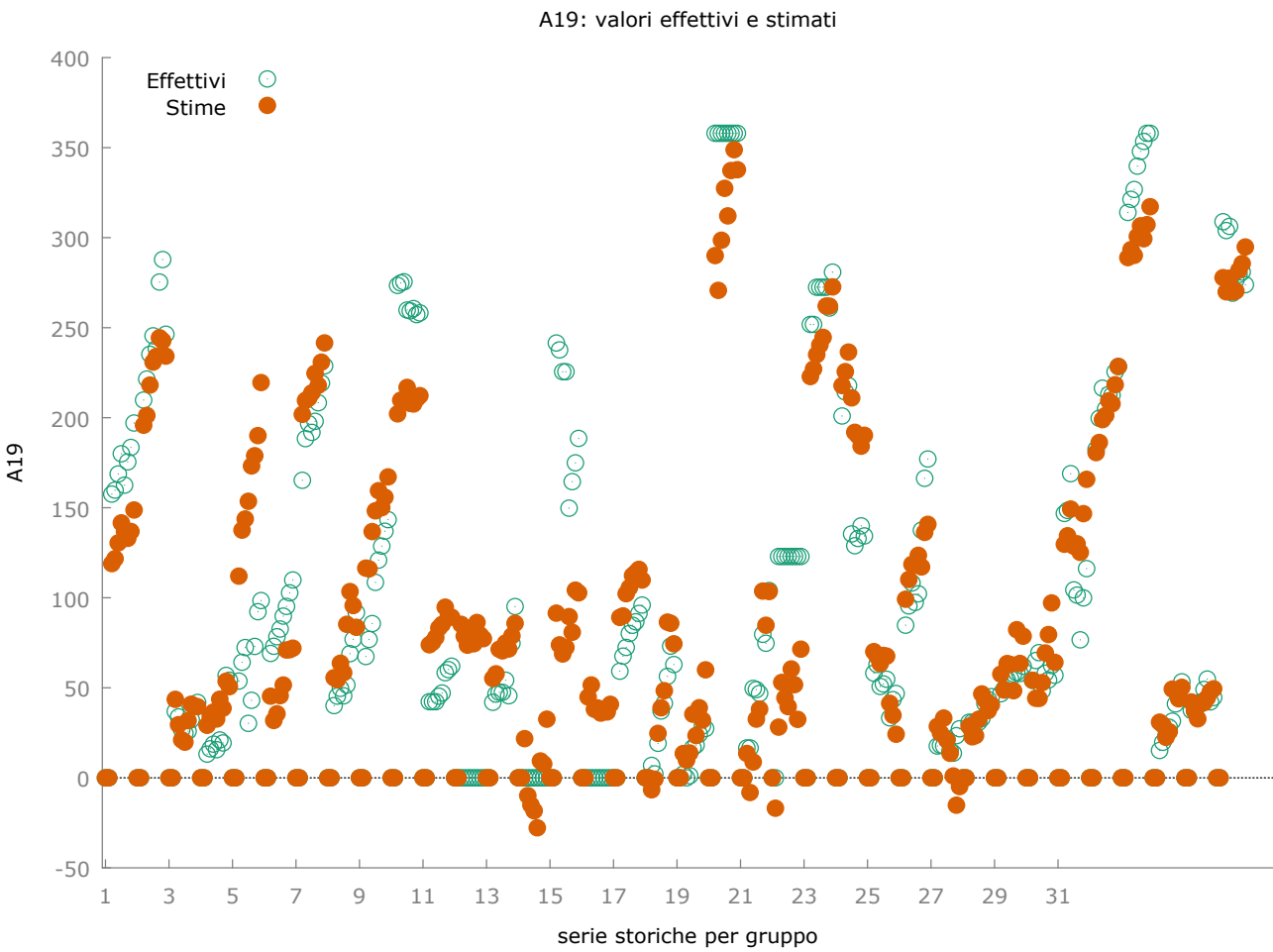
| | | | | | |
|-----------------------|-----------|------------------------|----------|---------|-----|
| A56 | -0,259626 | 0,0465016 | -5,583 | <0,0001 | *** |
| Media var. dipendente | 89,82317 | SQM var. dipendente | 101,2438 | | |
| Somma quadr. residui | 497020,8 | E.S. della regressione | 37,84622 | | |
| Log-verosimiglianza | -1812,269 | Criterio di Akaike | 3652,538 | | |
| Criterio di Schwarz | 3706,943 | Hannan-Quinn | 3674,170 | | |
| rho | 0,486969 | Durbin-Watson | 0,788003 | | |

Varianza 'between' = 1162,55
Varianza 'within' = 493,11
Theta usato per la trasformazione = 0,798282

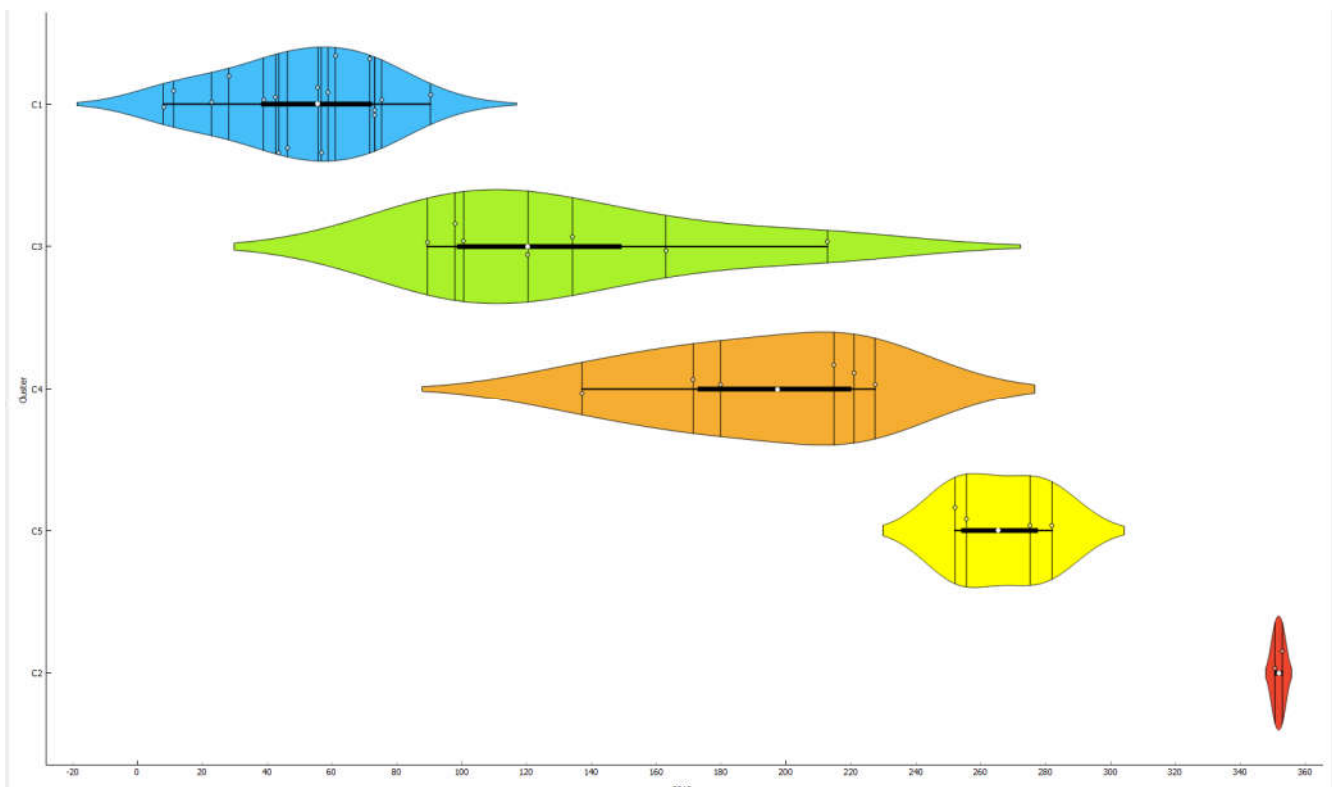
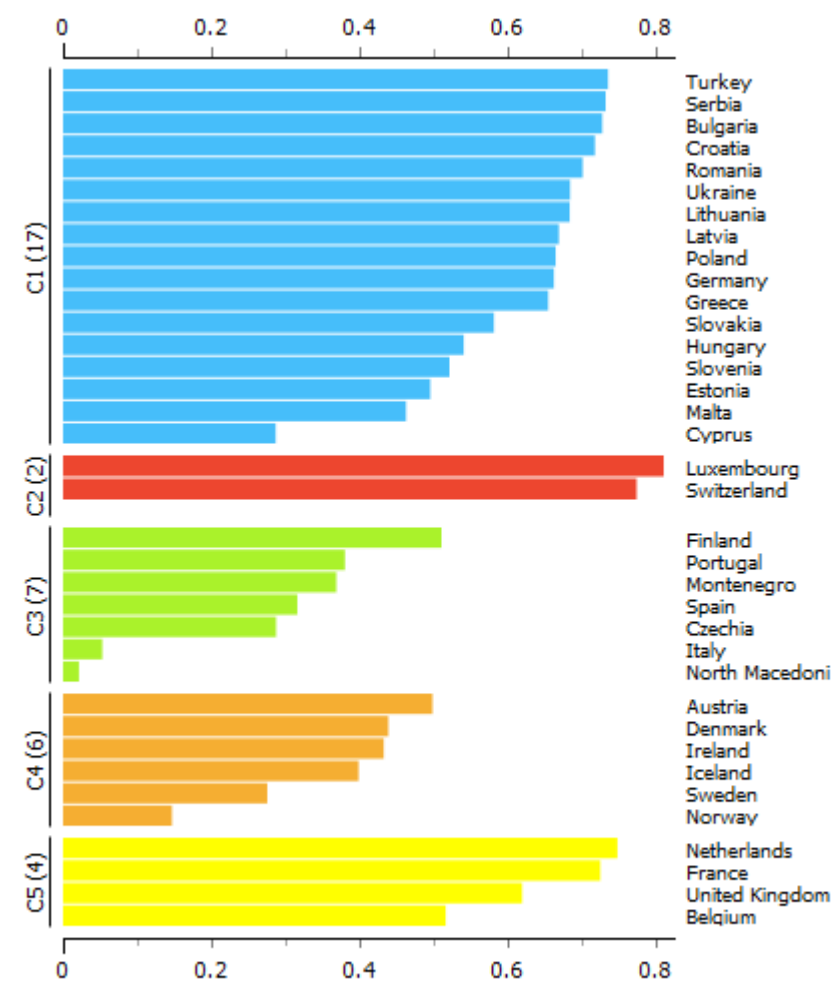
Test congiunto sui regressori -
Statistica test asintotica: Chi-quadro(13) = 2708,69
con p-value = 0

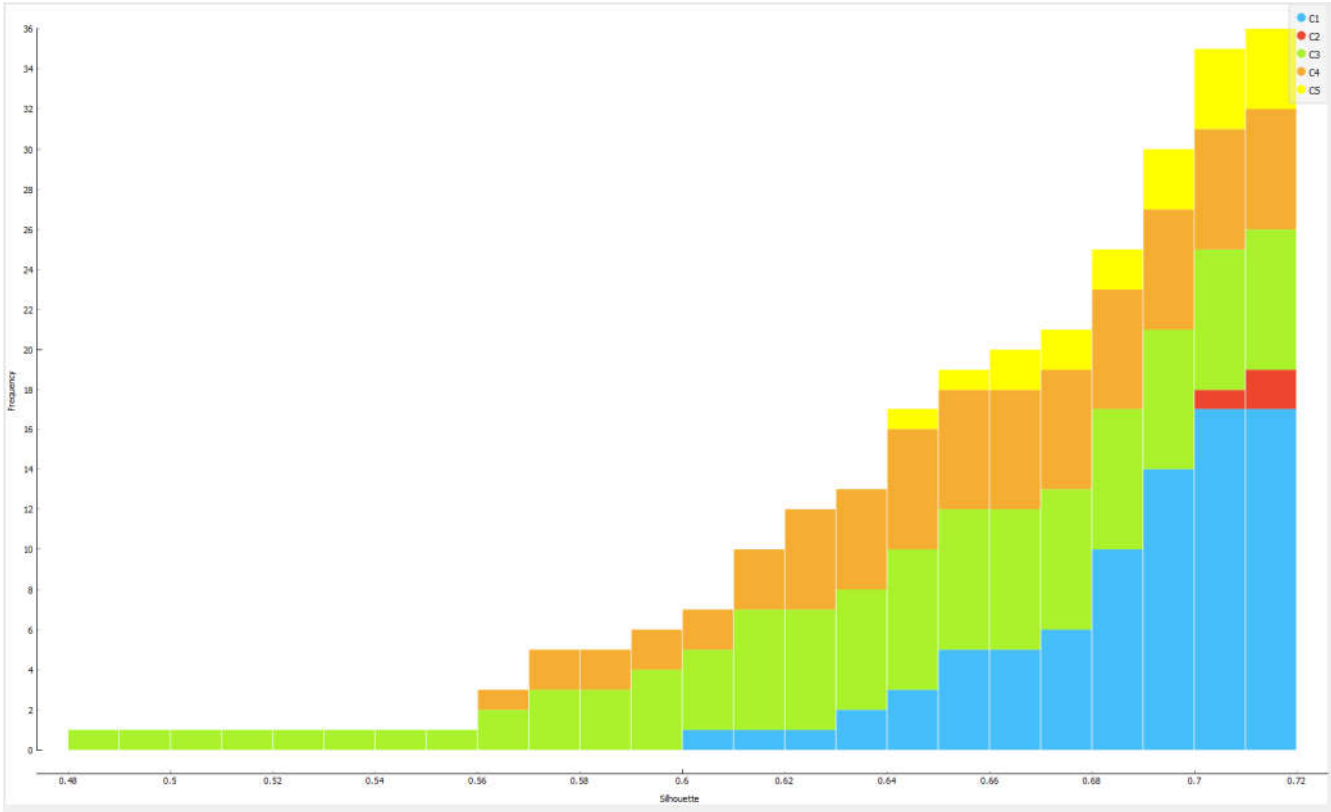
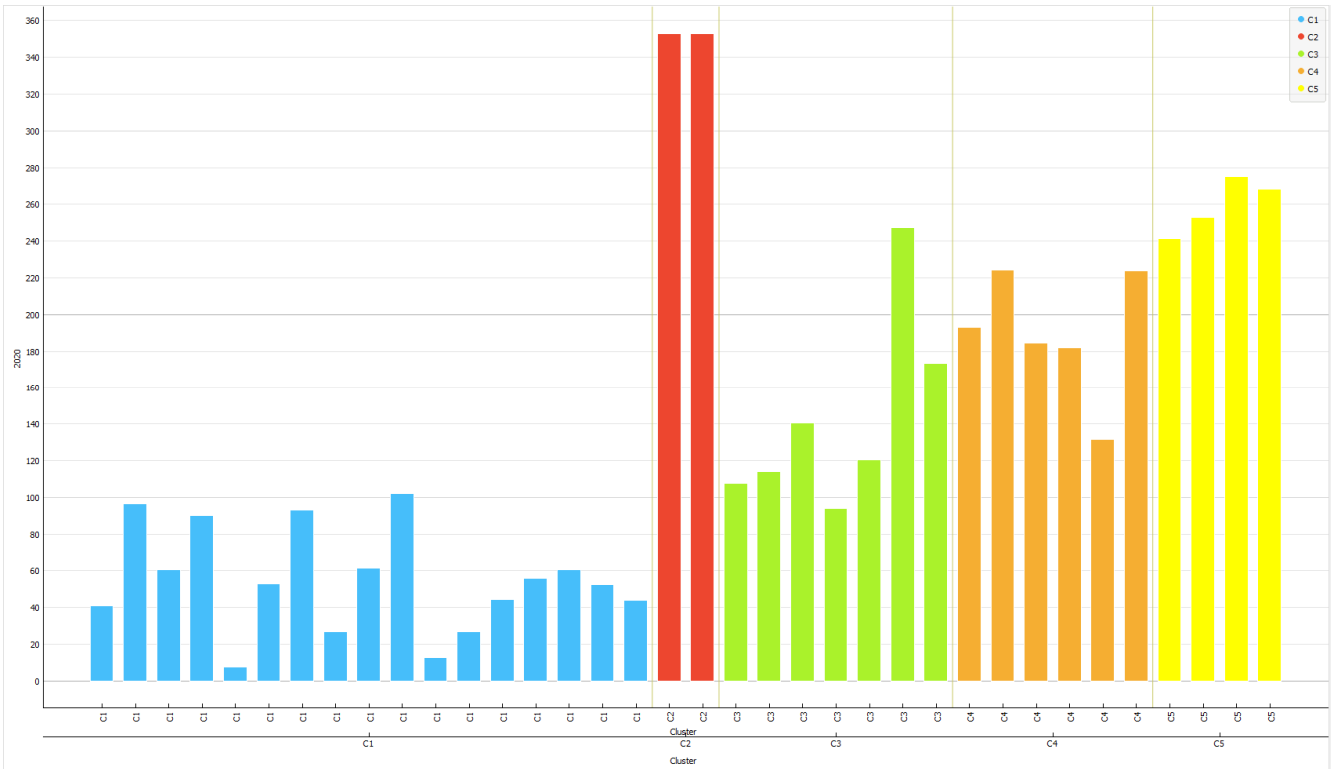
Test Breusch-Pagan -
Ipotesi nulla: varianza dell'errore specifico all'unità = 0
Statistica test asintotica: Chi-quadro(1) = 537,023
con p-value = 8,37546e-119

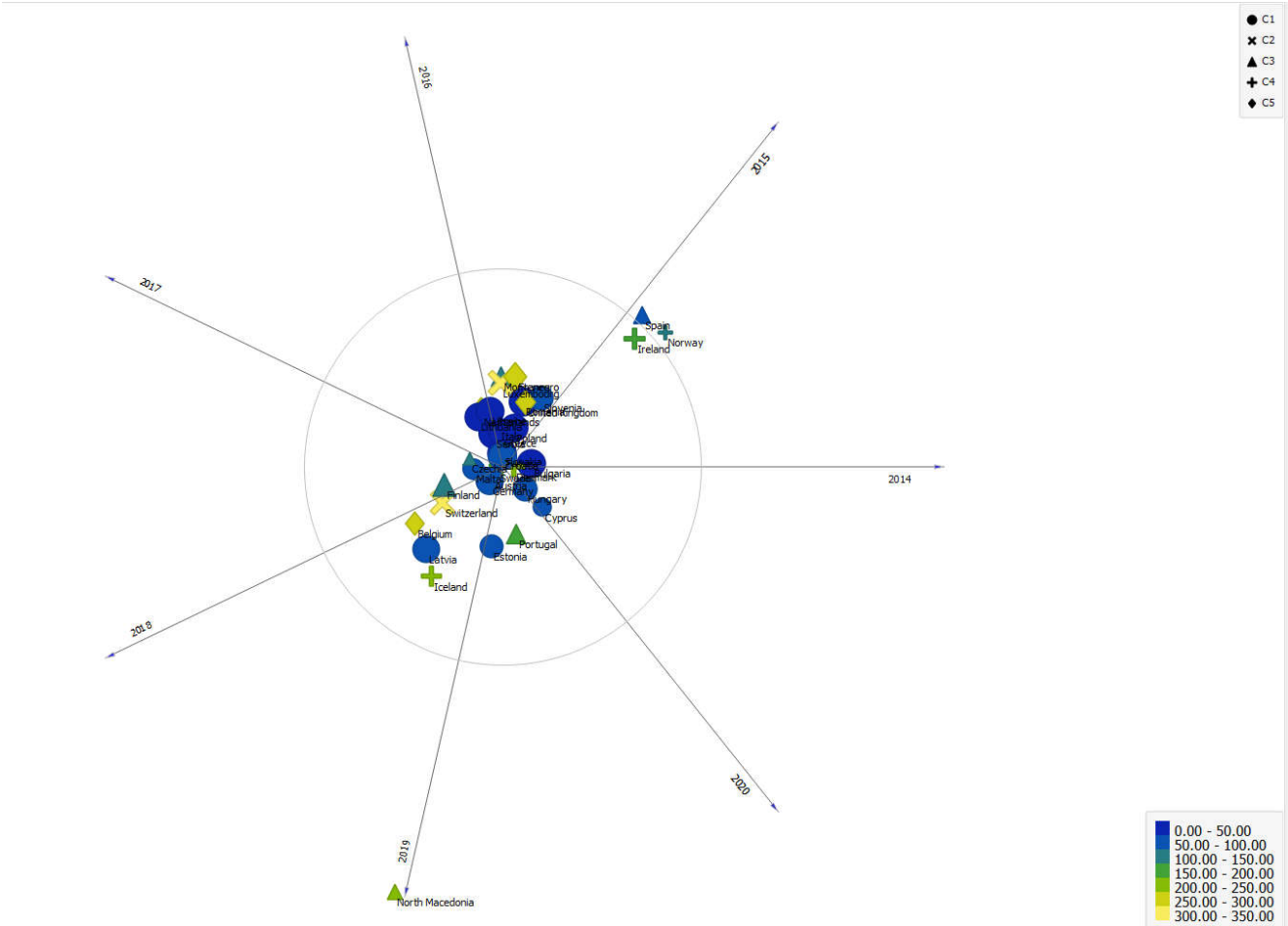
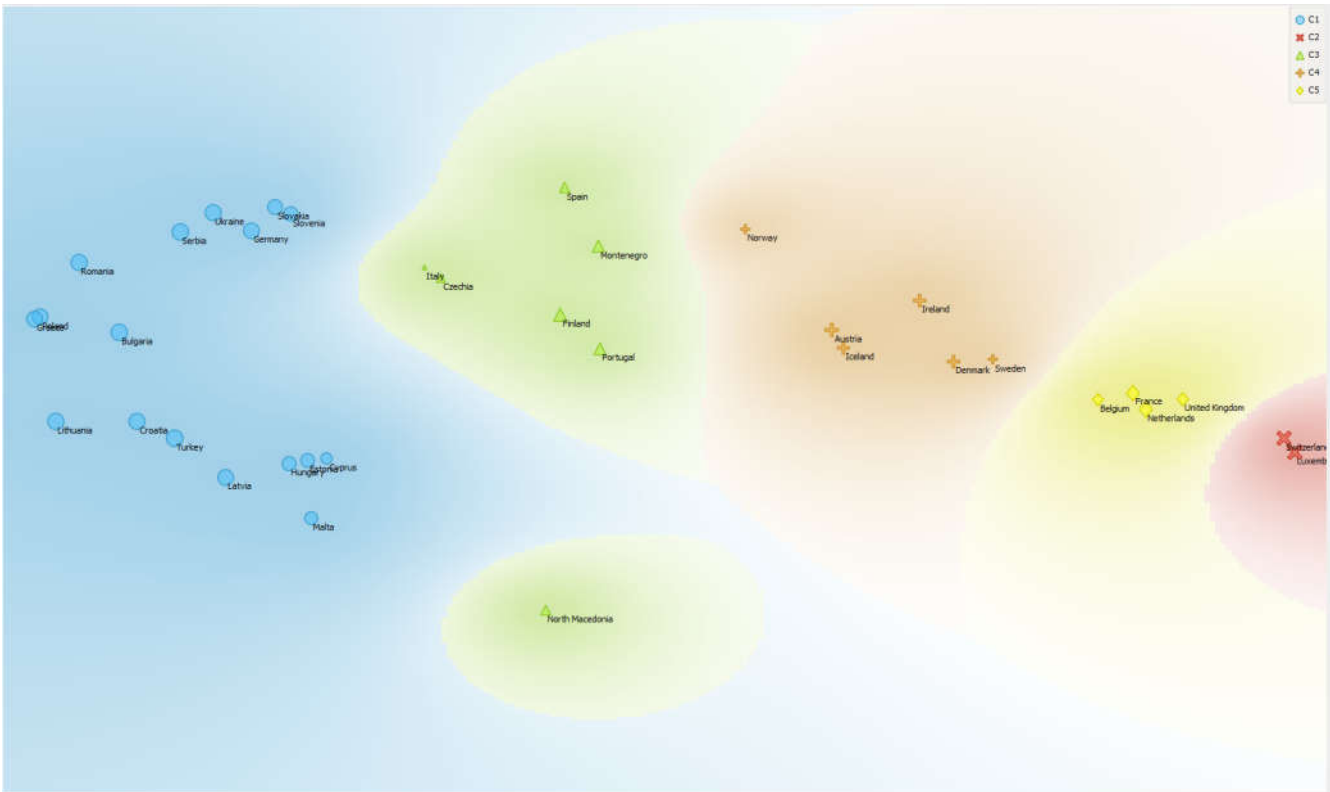
Test di Hausman -
Ipotesi nulla: le stime GLS sono consistenti
Statistica test asintotica: Chi-quadro(13) = 7,34339
con p-value = 0,883599

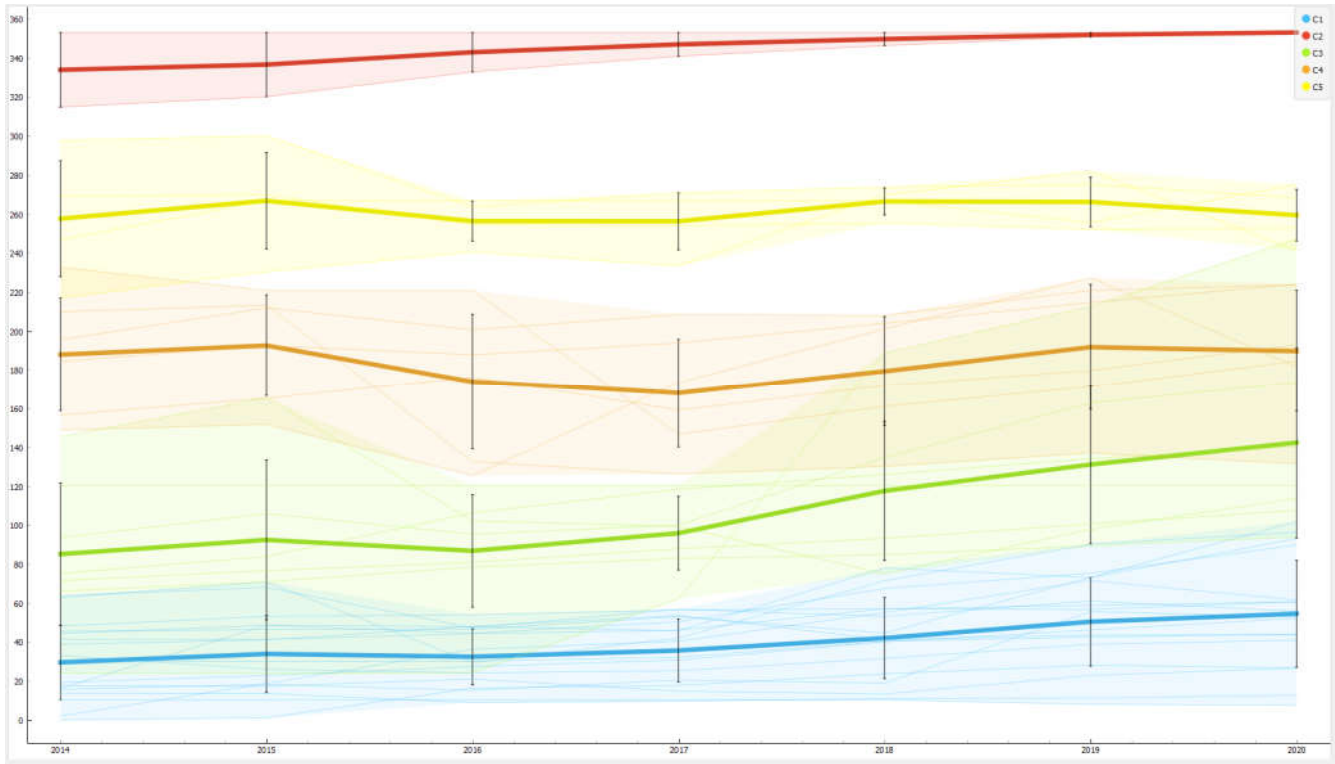
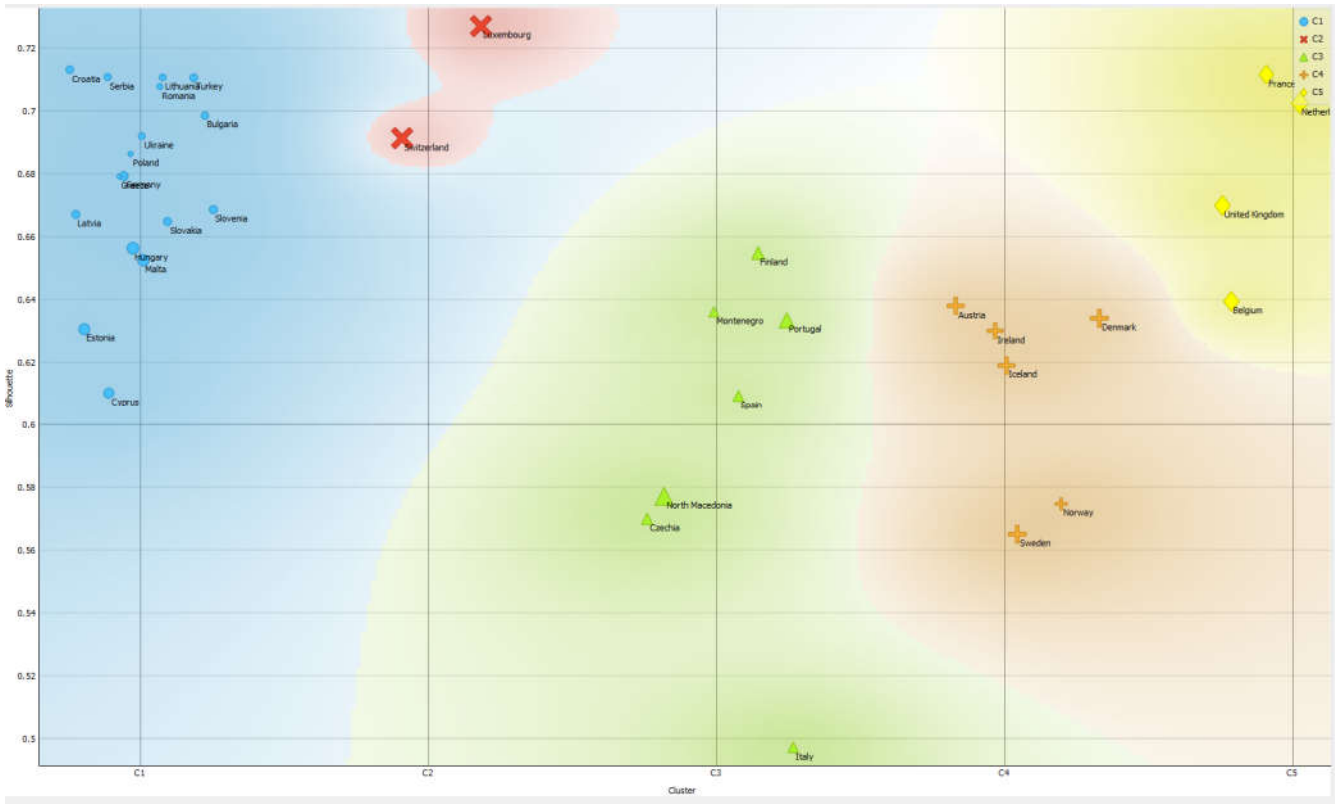


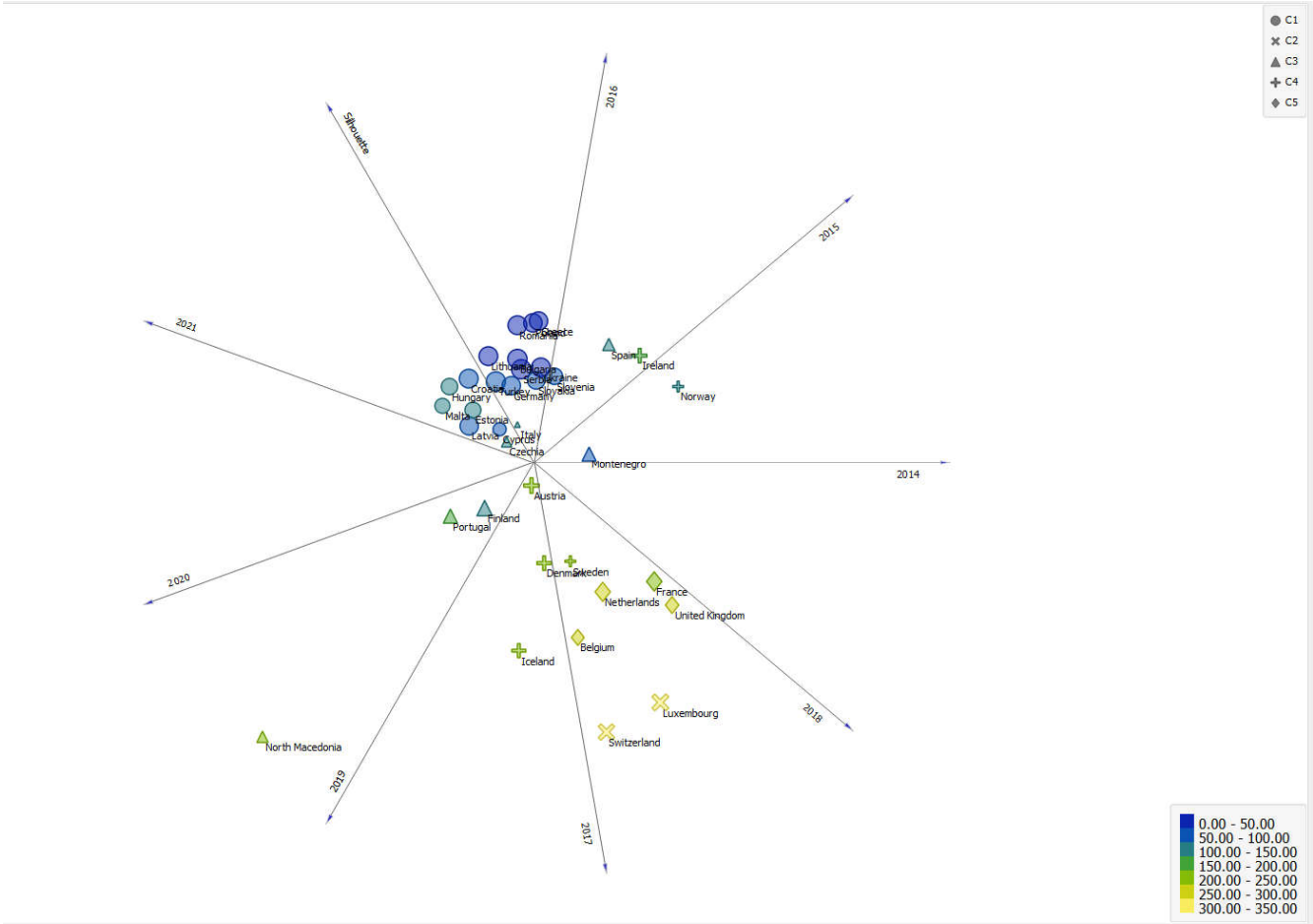
Clusterization

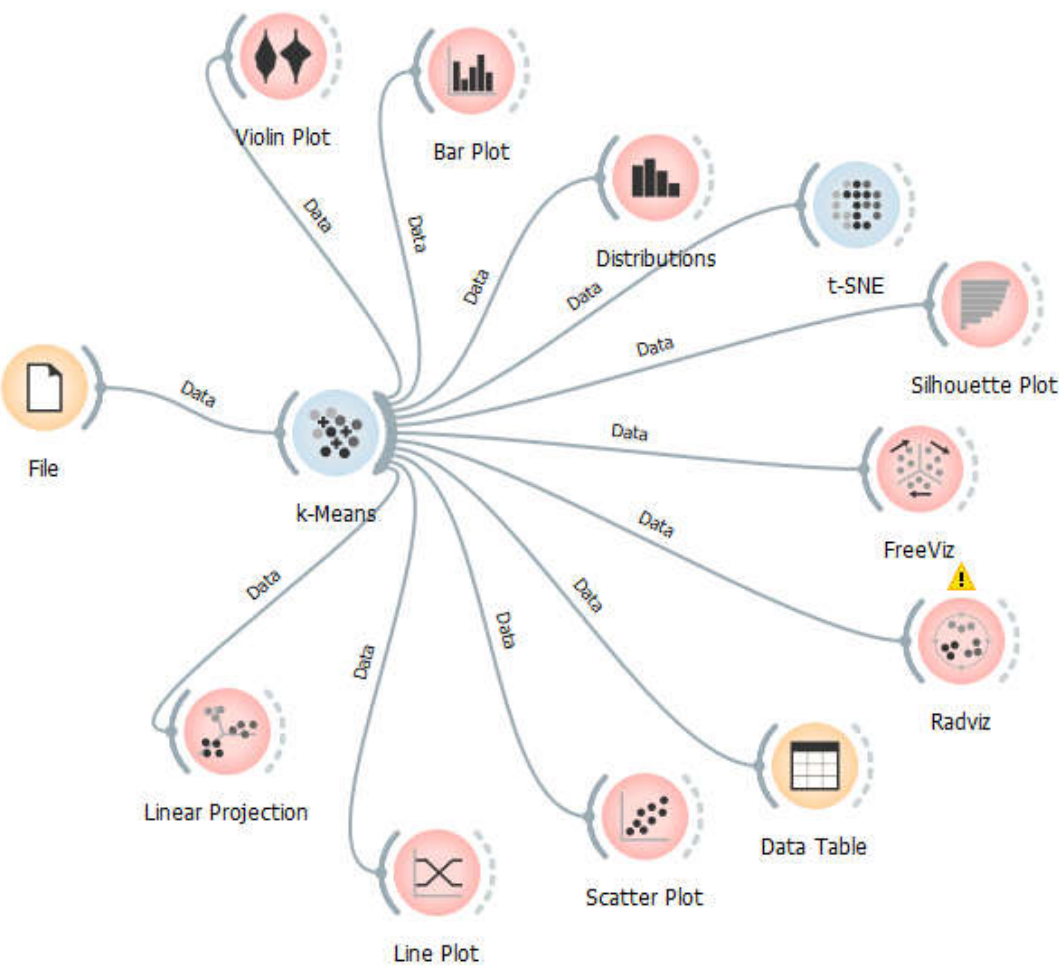






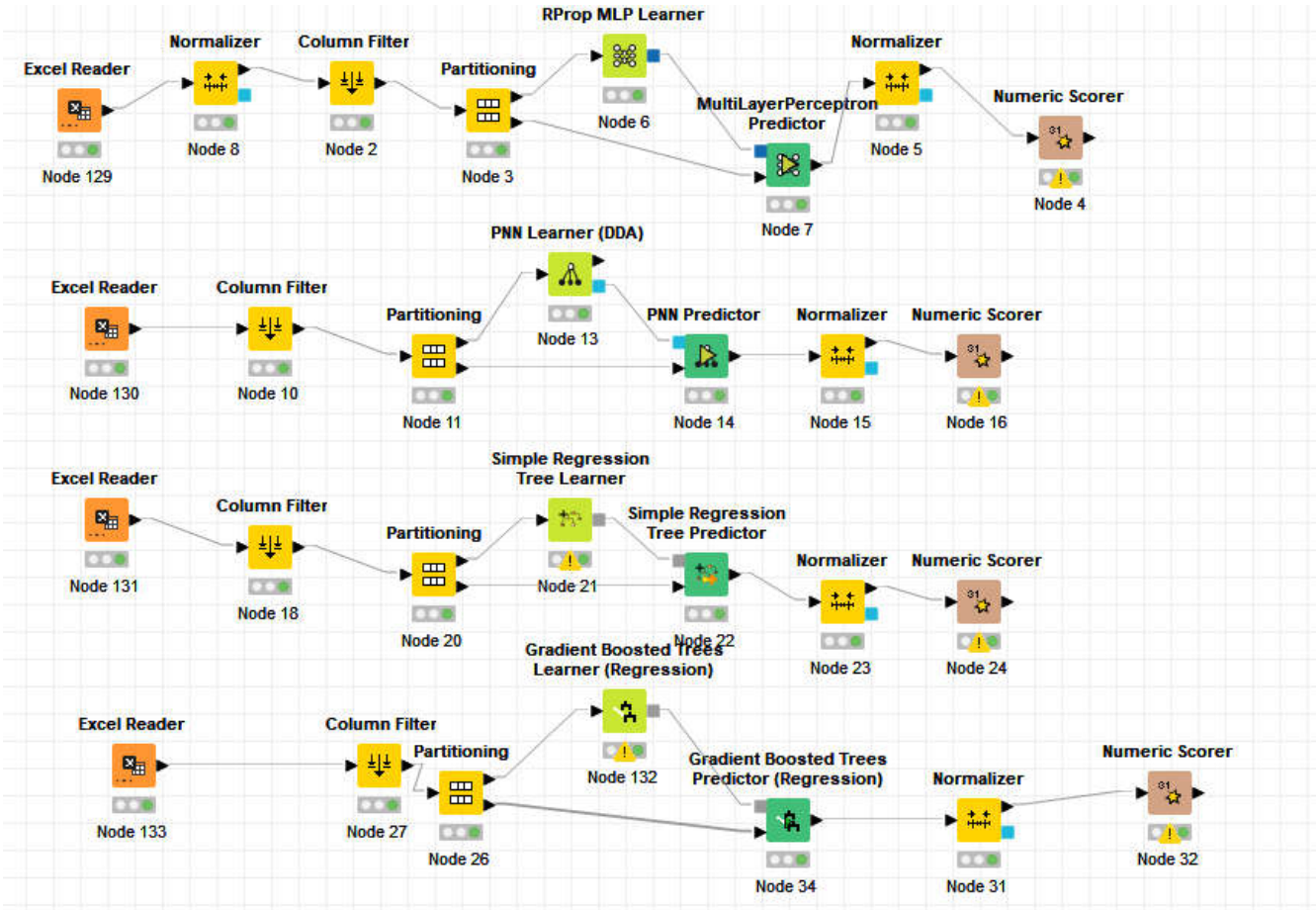


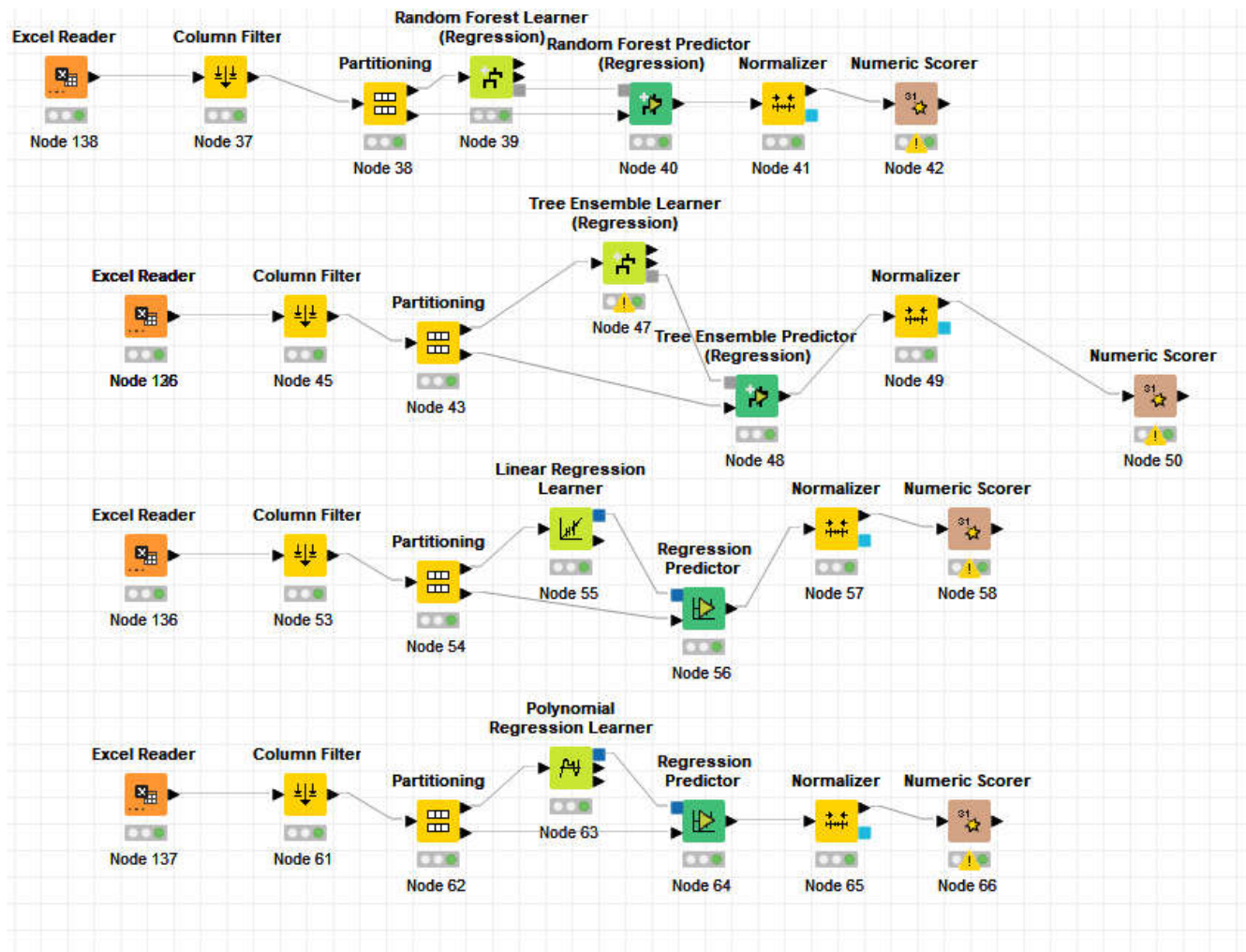




| | 2021 | Country | Cluster | Silhouette |
|----|--------|-----------------|---------|------------|
| 3 | 46.26 | Bulgaria | C1 | 0.703124 |
| 5 | 93.62 | Cyprus | C1 | 0.602704 |
| 7 | 75.14 | Germany | C1 | 0.688506 |
| 9 | 109.51 | Estonia | C1 | 0.649249 |
| 10 | 6.87 | Greece | C1 | 0.686491 |
| 14 | 57.90 | Croatia | C1 | 0.697327 |
| 15 | 123.03 | Hungary | C1 | 0.656952 |
| 19 | 42.40 | Lithuania | C1 | 0.693635 |
| 21 | 63.20 | Latvia | C1 | 0.684926 |
| 24 | 120.78 | Malta | C1 | 0.634744 |
| 27 | 13.47 | Poland | C1 | 0.689027 |
| 29 | 28.45 | Romania | C1 | 0.696522 |
| 30 | 45.87 | Serbia | C1 | 0.703814 |
| 32 | 61.57 | Slovenia | C1 | 0.65855 |
| 33 | 61.32 | Slovakia | C1 | 0.673505 |
| 34 | 52.82 | Turkey | C1 | 0.703341 |
| 35 | 43.71 | Ukraine | C1 | 0.691932 |
| 4 | 353.09 | Switzerland | C2 | 0.702341 |
| 20 | 353.09 | Luxembourg | C2 | 0.710335 |
| 6 | 116.80 | Czechia | C3 | 0.564076 |
| 11 | 108.63 | Spain | C3 | 0.591866 |
| 12 | 143.92 | Finland | C3 | 0.64614 |
| 18 | 102.11 | Italy | C3 | 0.48581 |
| 22 | 99.63 | Montenegro | C3 | 0.618002 |
| 23 | 234.66 | North Macedonia | C3 | 0.57483 |
| 28 | 184.87 | Portugal | C3 | 0.617152 |
| 1 | 212.00 | Austria | C4 | 0.645758 |
| 8 | 230.97 | Denmark | C4 | 0.626602 |
| 16 | 193.07 | Ireland | C4 | 0.629164 |
| 17 | 213.01 | Iceland | C4 | 0.618954 |
| 26 | 133.72 | Norway | C4 | 0.56538 |
| 31 | 226.51 | Sweden | C4 | 0.57036 |
| 2 | 264.84 | Belgium | C5 | 0.64995 |
| 13 | 243.69 | France | C5 | 0.699684 |
| 25 | 280.86 | Netherlands | C5 | 0.700058 |
| 36 | 264.63 | United Kingdom | C5 | 0.667143 |

Machine learning and predictions





References

- Amiri, F. & Puteh, M., 2018. Oral Communication Apprehension among International Doctoral Students. In: *English Language Teaching*. 11: 2, pp. 164-171.
- Baloch, N., Siming, L., Abraha, A. & Hong, S., 2021. Faculty research productivity: differences between foreign and local doctoral degree holders in Pakistan. *Higher Education*, 82(1), pp. 203-225.
- Che Arr, F., 2021. South East Asian Female Doctoral Students' Sojourning Experience in New Zealand: A Process of Gaining by Losing. *Doctoral dissertation*, Volume Auckland University of Technology.
- Doyle, S. et al., 2018. African international doctoral students in New Zealand: Englishes, doctoral writing and intercultural supervision. *Higher Education Research & Development*, 37(1), pp. 1-14.
- Fenyves, V. et al., 2019. Analysis of factors influencing foreign studies—strategic decisions—results of a Hungarian survey. *Journal of Entrepreneurship Education*, 22(5), pp. 1-21.
- Hou, M. & Jam, A., 2020. International Education Equity for Doctoral Students: Duoethnographic Reflections from China and Cameroon. *International Journal of Doctoral Studies*, 15(759).
- Interiano, C. G. & Lim, J. H., 2018. A "chameleonic" identity: Foreign-born doctoral students in US counselor education. *International Journal for the Advancement of Counselling*, 40(3), pp. 310-325.
- Jung, J., 2019. Local and Non-local Doctoral Students in Hong Kong: Do Stressors Differ with Students' Origins?. *International Journal of Chinese Education*, 8(2), pp. 160-185.
- Kahn, M., Gamedze, T. & Oghenetega, J., 2019. Mobility of sub-Saharan Africa doctoral graduates from South African universities—A tracer study. *International Journal of Educational Development*, Volume 68, pp. 9-14.
- Lee, E., 2020. Re-thinking Engineering Doctoral Students' Sense of Belonging: In Consideration of Diversity in Citizenship and Interpersonal Interactions. *Doctoral dissertation*, Arizona State University.
- Lee, M. C., McMahon, M. & Watson, M., 2018. Career decisions of international Chinese doctoral students: The influence of the self in the environment. *Australian Journal of Career Development*, 27(11), pp. 29-39.
- Li, F. S. & Qi, H., 2019. An investigation of push and pull motivations of Chinese tourism doctoral students studying overseas. *Journal of Hospitality, Leisure, Sport & Tourism Education*, Volume 24, pp. 90-99.

-
- Lim, H. & Shin, J. C., 2018. Doctoral education in South Korea: On the way toward becoming an independent research hub. In: *Doctoral education for the knowledge society*. Cham: Springer, pp. 183-202.
- Lopes, B. & Lourenço, M., 2019. Unveiling 'European' and 'International' researcher identities: A case study with doctoral students in the humanities and social sciences. *Social Sciences*, 8(11), p. 303.
- Mao, J., 2022. Practicing the Rhythms of an International Doctoral Student's Life through Compassion, Connection, Commitment and Creativity: Enactment of Learner Agency. *Journal of International Students*, 12(4).
- Mathies, C. K. H., 2021. Do they stay or go? Analysis of international students in Finland. *Globalisation, Societies and Education*, 19(3), pp. 298-310.
- Nguyen, L., 2018. Becoming professional researchers: An exploration of the experiences of Vietnamese doctoral students in Australia. *Doctoral dissertation*.
- Nomnian, S., 2018. Thai doctoral students' layers of identity options through social acculturation in Australia. *ASEAS-Austrian Journal of South-East Asian Studies*, 11(1), pp. 99-116.
- Pappa, S., Elomaa, M. & Perälä-Littunen, S., 2020. Sources of stress and scholarly identity: the case of international doctoral students of education in Finland. *Higher Education*, 80(1), pp. 173-192.
- Todoran, C. & Peterson, C., 2020. Should they stay or should they go? How the 2017 US travel ban affects international doctoral students. *Journal of Studies in International Education*, 24(4), pp. 440-455.
- Wang, L. & Byram, M., 2019. International doctoral students' experience of supervision: a case study in a Chinese university. *Cambridge Journal of Education*, 49(3), pp. 255-274.
- Wang, Q., Wang, H. H. & Zhang, J., 2020. Contributions of Chinese doctoral recipients from foreign universities to agricultural economic research and education in North America and China. *China Agricultural Economic Review*.
- Xu, L. & Hu, J., 2020. Language feedback responses, voices and identity (re) construction: Experiences of Chinese international doctoral students. *Innovations in Education and Teaching International*, 57(6), pp. 724-735.
- Xu, L. & Zhang, L. J., 2019. L2 doctoral students' experiences in thesis writing in an English-medium university in New Zealand. *Journal of English for Academic Purposes*, 41(100779).
- Yang, Y., Volet, S. & Mansfield, C., 2018. Motivations and influences in Chinese international doctoral students' decision for STEM study abroad. *Educational Studies*, 44(3), pp. 264-278.